Air Force Institute of Technology
AFIT Scholar

Theses and Dissertations | Student Graduate Works

3-5-2004

Developing a Forecast Tool for Cloud-to-Ground Lightning in the North Central and Northeastern United States

Manuel I. Folsom Jr.

DEVELOPING A FORECAST TOOL FOR CLOUD-TO-GROUND LIGHTNING IN THE NORTH CENTRAL AND NORTHEAST UNITED STATES

THESIS

Manuel I. Folsom Jr., Captain, USAF

AFIT/GM/ENP/04-05

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GM/ENP/04-05

DEVELOPING A FORECAST TOOL FOR CLOUD-TO-GROUND LIGHTNING

IN THE NORTH CENTRAL AND NORTHEAST UNITED STATES

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Meteorology

Manuel I. Folsom Jr., BS

Captain, USAF

March 2004

AFIT/GM/ENP/04-05

DEVELOPING A FORECAST TOOL FOR CLOUD-TO-GROUND LIGHTNING IN THE NORTH CENTRAL AND NORTHEAST UNITED STATES

Manuel I. Folsom Jr., BS
Captain, USAF

Approved:

Michael K. Walters (Chairman)	5 MAR 04 date
Ronald P. Lowther (Member)	5 MAR 04 date
Robin N. Benton (Member)	5 Mar 04 date

AFIT/GM/ENP/04-05

Abstract

Cloud-to-ground (CG) lightning is a hazard to the Air Force for both air and ground operations. Forecasting CG lightning is a necessary and extremely important requirement for Air Force meteorologists and forecasters. The 15th Operational Weather Squadron requested a forecast tool capable of predicting CG lightning within a 25 and 10 nautical mile radius of the 13 military locations in their area of responsibility. To fulfill their request, forecast decision tools were created using a Classification and Regression Tree (CART) data analysis program.

Four decision trees were produced for each location using the period of record from March through September, 1993 to 2002. CART compared the upper air stability indices and surface data at 12-hour intervals with CG lightning data occurring within the next 12 hours to determine prediction rules. Data from 2003 were used as independent verification of the decision trees.

The CART decisions trees were examined using contingency tables and verification tests to determine the value of the products created. The straight forward forecast rules and verification test results confirmed that the decision trees would be a valuable prediction tool. Combined with forecaster knowledge, forecast models and other tools, the decision trees would provide an excellent forecast method for determining the occurrence of CG lightning. Therefore, the results are recommended to the 15th Operational Weather Squadron for use as a CG lightning forecast tool.

Acknowledgments

I would like to express my sincere appreciation to my faculty advisor, Lt Col Michael Walters, for his guidance and support throughout the course of this thesis effort. The insight and experience was certainly appreciated. I would also like to thank my committee, Lt Col Ronald Lowther and Maj Robin Benton for providing additional knowledge toward the completion of this endeavor.

I would like to thank Mr. Jeffrey Zautner and Mr. Pete Speck from the Air Force Climatology Center for providing the data. I would also like to thank my sponsor, the 15th Operational Weather Squadron for providing information for the Northeast and Northern Midwest regions.

I would like to thank my classmates, especially Jonathan Leffler and Steve Callis for their excellent ideas and generous help towards the completion of this thesis.

Finally, I am indebted to my family for their love and support throughout my entire effort to accomplish this task.

Manuel I Folsom Jr.

Table of Contents

Page

List of Figures

List of Tables

DEVELOPING A FORECAST TOOL FOR CLOUD-TO-GROUND LIGHTNING IN THE NORTH CENTRAL AND NORTHEAST UNITED STATES

1. Introduction

Thunderstorms are weather phenomena that affect the entire region of North America. A major product of these thunderstorms, cloud-to-ground (CG) lightning is a major hazard to the weather and Air Force communities. CG Lightning poses the greatest threat to people and property during the thunderstorm season (NWS 1997). The following statistics were reported in a National Oceanic and Atmospheric Administration (NOAA) Technical Memorandum of thunderstorm and CG lightning related incidents reported from 1959-1994 (NWS 1997). There were 3,229 deaths, 9,818 injuries, and 19,814 reports of property damage in the United States. Young men of age 10-35, either at work or involved in recreation, and children accounted for 85 percent of lightning victims. The highest density of these incidents occurred in New England, the southeast, the plains, and the Rocky Mountains (NWS 1997).

The world's increasing dependence on technology and energy has increased the importance for the weather community to achieve the ability to predict the occurrence of thunderstorms and CG lightning to help minimize the damage to electrical supply lines and to system technologies. The commercial industry reported estimates totaling nearly

This thesis follows the style of the American Meteorology Society (AMS 2002).

$385 million dollars in damage structures and electrical systems from 1973-1982 (NLSI 2003). CG lightning can be blamed for most of the claims since it is the single largest cause of outages in electric power distribution systems (NLSI 2003).

The Department of Defense (DoD) and the Federal Aviation Administration (FAA) are extremely concerned with CG lightning due to the risk to personnel and the cost of replacing damaged equipment. With a single aircraft costing millions of dollars, the need for better thunderstorm and lightning warnings is becoming more crucial than ever. The 15th Operational Weather Squadron (OWS) located at Scott AFB, Illinois has the responsibility of forecasting weather for two of these critical regions. The 15th OWS issues forecasts, watches, and warnings for 12 active duty and 17 Reserve and National Guard military installations across the region shown in Fig. 1. From March through September 2003, the 15th OWS verified 747 separate cases of CG lightning in their area of responsibilities (AOR).

FIG 1. The 15th Operational Weather Squadron Area of Responsibility (AOR) (modified from 15th OWS 2003).

a. Statement of the Problem

The 15th OWS has a need for a reliable forecast tool to predict CG lightning. Venzke (2001) developed a method to apply stability indices computed from upper air soundings to determine the probability of CG lightning for a region within a 50 nautical-mile (nm) radius surrounding the upper air stations. This research started with the same problem but will approach the problem using different prediction tools and prediction parameters. The focus will be placed on a different region of the U.S. and this study will use different location criteria in determining the locations used. Venzke (2001) developed predictive tools for CG lightning that could be applied to the Midwest but for only locations with an upper air site on location. The area of interest for this research will be in the North Central and Northeast U.S., not the Midwest, and the focus will be on locations geographically separated from the upper air stations, in some cases by over 100 miles. Forecasters at most military installations do not have the benefit of on-site upper air stations, therefore this research will focus on using what is available to the military operational forecasters in an effort to determine a predictor for CG lightning for the radial distance that is required at each location. The National Weather Service (NWS) report NWS SR-193 (1997) indicated that the critical time frame for CG lightning for the North Central and Northeast was centered on the summer months during the afternoon hours, therefore this study will focus on a data set containing the warmer months.

The severity of the thunderstorms was not considered in this research. This research was only concerned with the occurrence of CG lightning. Frontal systems are a major concern when using data from two geographically separated sources. The greatest potential shortfall of using stability indices alone is the possible difference between the air masses at the upper air site and the point of interest. Therefore, surface data from both the upper air location and the location of interest were compared and implemented into the prediction process to reduce the effects of frontal systems on the decision tree.

b. Research Objectives

One of the goals for this research was to find a good, user friendly prediction tool for CG lightning for each of the locations in the 15th OWS AOR. Keeping this main objective in mind, the data criterion was to limit the predictors to a list that could easily be created and used by the forecaster in the field. Another objective was to use the prediction tool to forecast the probability of lightning within two categories: 25 nautical miles (nm) and 10 nm. The following tasks were required to achieve the forecast products for each of the North Central and Northeast US locations:

1. Communicate with the 15^{th} OWS to determine the upper air locations best suited for use in each of the 13 locations within their AOR.
2. Focus on the most critical time of year for the CG lightning threat and choose the predictors to be used in the data set.
3. Use Classification and Regression Tree (CART) analysis to determine if predictive tools can be formulated for each location of interest.

4. Apply CART to the data set containing CG lightning counts and stability indices to determine a location dependent list of rules for the occurrence of CG lightning.
5. Create decision tools based on the findings from the CART output.
6. Use an independent data set to validate the decision tools.

This research combined the knowledge learned from both Venzke (2001) and Sahu and Singh (1999) to determine the best methods for developing the CG lightning forecast decision trees. CART, a new and exciting statistical program, was applied to the process performed by Venzke (2001) to develop a better tool for forecasting CG lightning. Ideas for new predictors were taken from the works of Sahu and Singh (1999). Data sets from March 1993 through September 2002 were analyzed using CART to develop decision trees for the prediction of CG lightning. The decision trees were validated using independent data from March through September 2003.

Chapter 2 provides a background discussion for the study. Previous studies on lightning and stability indices were used to discover the important findings from other authors. Chapter 3 focuses on the sources of the data sets and the methodology used to formulate the research results. Chapter 4 discusses the results found using the CART analysis program. Decision tools and forecast rules are provided for future use by the 15th OWS. The final chapter summarizes the research and discusses the importance of the results to the field of meteorology and to the mission of the 15th OWS.

2. Literature Review

a. Lightning

Thunderstorms are of major concern to the aviation and business communities. Lightning discharges can occur as inter-cloud, cloud-to-cloud, cloud-to-air, and cloud-to-ground. A top weather concern, CG lightning can kill, destroy equipment, start fires, and disturb power deliveries systems. The National Oceanic and Atmospheric Administration (NOAA) performed a study on the effects of CG lightning across the United States. The NOAA Technical Memorandum (NWS 1997) summarized damage, injuries, and deaths resulting from lightning in the United States from 1959-1994. The variability of lightning-caused reports was less than any other weather event during this period; therefore, people are vulnerable to CG lightning throughout the entire thunderstorm season. People involved in recreation, located around trees, or located close to bodies of water were the three largest categories of lightning fatalities. A statistical comparison of all weather phenomena lists CG lightning as the second leading cause of weather related deaths behind the combined category of flash and river flood related deaths. Table 1 shows the top ten locations for deaths, injuries, and damage reported for the period 1959-1994.

Table 1. Top ten states for lightning caused deaths, injuries, and damage. The states in bold are located within the 15th OWS AOR (modified from NWS SR-193 1997).

Rank	Deaths	Number of Deaths	Injuries	Number of Injuries	Damage	Number of Damage Reports
1	Florida	345	Florida	1178	**Pennsylvania**	1441
2	North Carolina	165	**Michigan**	643	Kansas	1182
3	Texas	164	**Pennsylvania**	535	**New York**	1005
4	**New York**	128	North Carolina	464	North Carolina	960
5	Tennessee	124	**New York**	449	Oklahoma	826
6	Louisiana	116	**Ohio**	430	**Michigan**	814
7	**Maryland**	116	Tennessee	349	Tennessee	764
8	**Ohio**	115	Texas	334	South Carolina	717
9	Arkansas	110	**Massachusetts**	331	Texas	689
10	**Pennsylvania**	109	Georgia	329	Georgia	656

1) NATIONAL LIGHTNING DETECTION NETWORK (NLDN)

Forecasters verified thunderstorms through three main processes: 1) using satellite data, which were thirty minutes or older, 2) using radar data, which were several minutes to thirty minutes old, and 3) using the weather observations which were station dependent, so if thunderstorms were not in the vicinity of a reporting station, they were not reported. The timeliness of data available to weather forecasters demonstrated the need of a near time reporting system for thunderstorms. A device to detect CG lightning would also be beneficial for issuing thunderstorm warnings to protect people at large outdoor events or engaged in sporting activities. The National Lightning Detection Network (NLDN) was created in 1988 to provide near real time data and archived data to those with a concern for CG lightning. The location accuracy of the system was 8 to 16 km with a first stroke flash detection efficiency of 70%. In 1992, GeoMet Data Services (GDS), the company that maintained the NLDN, calibrated the sensors to increase the location accuracy to 4 to 8 km, and the first stroke flash detection efficiency was also improved to 65 to 80%. A major upgrade to the lightning detection system was completed in 1995. This upgrade improved the location accuracy to 1 to 2 km and increased the first stroke flash detection efficiency to 80 to 90%. Video verification showed detection efficiencies at 84% prior to the upgrade in 1994 and 85% after the upgrade in 1995. However, the NLDN detects only cloud-to-ground lightning, therefore thunderstorm forecasts using this method were limited since a thunderstorm might only produce the undetectable intracloud lightning (Huffines and Orville 1999).

Wacker and Orville (1999a) analyzed the CG lightning data after the upgrade of 1994 to determine the benefits and results of the upgrade. The 1994 upgrade had three major goals: use the IMPACT (improved performance through combined technology) system standards to improve location accuracy, increase detection efficiency, and enhance the reliability of the NLDN hardware. Wacker and Orville (1999b) concluded that the number of detected weak flashes of CG lightning increased after the upgrade, therefore the total number of detected CG lightning increased. And, since the total number of strong flashes remained constant, the mean peak current actually decreased for detected flashes. A major result of the upgrade was that decreasing the pulse width detection criterion improved the NLDN but it also included an unwanted contamination of intra-cloud flashes (Wacker and Orville 1999b). A more complete discussion on the NLDN and the upgrade can be found in Wacker and Orville (1999 a,b) and Cummins et al. (1998).

Global Atmospherics Inc. (GAI), the company that currently maintains the NLDN, claims the system has a location accuracy of 500 meters and a detection probability between 80-90 percent, which varies slightly depending on region. One third of all lightning flashes contain strokes that strike the ground in different locations (GHRC 2003). Researchers defined a single flash as all discharges within 10 km and a one second interval (GHRC 2003).

The IMPACT system detects and groups all strokes from each flash. The lightning data contains the number of strokes per flash, time and location of flash, and peak signal amplitude. A major weather benefit of NLDN lightning data is that the strike locations can be determined in about 30 seconds. The data were available 24 hours a day, everyday of the year, and every lightning event is quality controlled to ensure accuracy.

2) APPLICATIONS OF NLDN DATA

The NLDN data set contains a continuous record of CG lightning across the continental United States extending 200-300 km off the coastline. These data provide a continuous data set dating from January 1, 1988 through the present. The NLDN data were used by National Weather Service, Federal Aviation Administration, The Weather Channel, PGA, major power companies, international and regional airports, and many businesses nationwide (Lightning Storm 2003).

The data also provided an excellent CG lightning research tool since, for the first time, total area coverage was available instead of station dependent data. Huffines and Orville (1999) calculated the lightning ground flash density and thunderstorm duration for the Continental United States. They demonstrated that for the period from 1989 thru 1996, the Midwest region and the Atlantic Seaboard are two of the three maximums for CG lightning mean annual flash rate. Orville and Huffines (2001) again summarized CG lightning, but this time they used data for the decade 1989-1998. They again observed a very high density of CG lightning throughout the Midwest. To show the true value of a total coverage lightning detection system, Orville et al. (2002), focusing on the years 1998-2000, expanded the CG lightning study to include Canada and termed the combined detection system the North American Lightning Detection Network (NALDN).The CG lightning densities from all three results had two maximums in the regions monitored by the 15th OWS.

b. Stability

As discussed earlier, lightning is very important to the Department of Defense and the civilian sector in terms of safety and property damage. The thunderstorms that produced this lightning required three critical elements for development: moisture, a trigger mechanism, and an atmosphere that is conditionally unstable. This section will focus on the need for a conditionally unstable atmosphere for thunderstorm development. Stability will be used as a prediction tool for determining the probability of lightning. The weather community uses several terms to describe the condition of the troposphere in terms of static stability: conditional instability, absolute instability (or stability) and convective instability. A brief description follows; however, for a more detailed discussion of stability, see Huschke (1959).

Static stability is a measure of the instability of the atmosphere and is a key requirement for thunderstorm development. Static stability is the measure of the stability of an atmosphere defined to be in hydrostatic equilibrium to vertical displacements (Huschke, 1959). Holton (1992) described static stability by using the dry adiabatic lapse rate and the atmospheric lapse rate. He stated the dry adiabatic lapse rate can be found in an atmosphere where potential temperature is constant with height. If the potential temperature increases with height, then the atmospheric lapse rate is less than the adiabatic lapse rate. The parcel, when adiabatically displaced from its equilibrium level, will be positively buoyant when displaced downward and will be negatively buoyant when displaced upward. The parcel will return to its original location to maintain hydrostatic equilibrium.

The stability of the atmosphere can be determined by comparing the lapse rate with the adiabatic lapse rate. The lapse rate is defined as the change in temperature with respect to pressure or height (Holton 1992). If the parcel is unsaturated, the comparison is made between the dry adiabatic lapse rate and the environmental lapse rate. The dry adiabatic lapse rate is calculated to be approximately 10°C km^{-1}.

Moisture in the atmosphere adds a complicating factor to the forecasting of the stability process. Conditional instability is defined as the of stability for the atmosphere in which the lapse rate is less than dry-adiabatic lapse rate but greater than the saturation adiabatic lapse rate (Wallace and Hobbs, 1977). The saturation adiabatic lapse rate is to be approximately 6.5°C km^{-1}. The atmosphere is unstable for the saturated parcel and the parcel if lifted adiabatically will displace vertically upward. However if the parcel is unsaturated, the adiabatically lifted parcel will return to its previous level.

Absolute stability and absolute instability are two terms used to describe the comparison between the lapse rate and the dry adiabatic lapse rate. An absolutely unstable atmosphere has a lapse rate greater than the dry adiabatic lapse rate. Therefore, if the parcel is lifted adiabatically, it will continue to accelerate upward. An absolutely stable atmosphere will have a lapse rate that is less than the dry adiabatic lapse rate and the saturated adiabatic lapse rate. Therefore if the parcel is lifted adiabatically, it will always return to its original level.

Convective instability, sometimes referred to as potential instability, is a term used to describe a column of air that has a wet bulb potential temperature and equivalent potential temperature that decreases with height. Theoretically, if this entire column of air is lifted upward, the top layer will cool faster than the lower layer. Therefore the lapse rate of the column will increase making the column unstable. Equivalent potential temperature is an important concept for thunderstorms in terms of downburst and hail potential because it is sensitive to the amount of moisture in the atmosphere (NWS 2003).

c. Stability Indices

Many stability indices have been developed by the weather community to describe the stability of the troposphere as a single numerical value to obtain more accurate thunderstorm and severe weather prediction. The strengths and weaknesses of each index are crucial for the forecaster to understand when applying them to the prediction of upcoming weather events. A critical point to keep in mind is if the indices are not utilized with other information, then the standard stability analyses of the complete sounding using the parcel method is a more useful method to use (AWS TR 79-006, 1990). The methods used to determine each index are very important to uncover the relevance of the index toward the weather pattern or conditions in the area of interest. An index that performs superbly on the Gulf Coast might be useless in the Midwest or Rocky Mountain regions.

Peppler (1998) described the original purpose for each of the stability indices. The Showalter Stability Index (SSI) and K Index (KI) were originally developed to predict non-severe thunderstorms and convective showers. The Lifted Index (LI), Vertical Totals (VT), Cross Totals (CT), Total Totals Index (TTI), and Severe Weather Threat Index (SWEAT) were originally created to forecast severe thunderstorms and tornadoes. Knowing how the stability index was developed and formulated provides valuable insight toward a better understanding of applying the index.

The subsequent sections contain brief descriptions of the indices used in this study, commonly used values, and an explanation of the atmospheric layers involved in the calculations of each index used in this research. The desired end state for this research is to determine if an index or a combination of indices would best predict CG lightning for the 13 locations under the watch of the 15th OWS at Scott AFB. A more detailed description of each index can be found in the articles published by the creators of the indices.

1) TOTAL TOTALS INDEX (TTI)

Miller (1972) originally developed the TTI to predict severe weather in areas where the potential for thunderstorm development was already forecasted through other means. However; since TTI is a commonly used index, it will be used to predict CG lightning for this research. TTI is a combination of two other commonly used indices, the Cross Totals (CT) and the Vertical Totals (VT). The standard units of TTI, CT, and VT are Celsius degrees. The TTI is computed using the following equations:

$$TTI = T(850\text{ mb}) + Td(850\text{ mb}) - 2[T(500\text{ mb})], \quad (1)$$

$$CT = Td(850\text{ mb}) - T(500\text{ mb}), \quad (2)$$

$$VT = T(850\text{ mb}) - T(500\text{ mb}). \quad (3)$$

Since CT is the difference between the dew point at 850 millibars (mb) and the temperature at 500 mb, CT combines critical low-level moisture information with the temperatures aloft. For regions east of the Rockies and along the Gulf Coast, CT is very effective for thunderstorm coverage and severity (AFWA 1998). However, if the moisture and cold air are not at the layers used in the CT formula, then CT will not be an effective indicator of thunderstorms. VT measures the lapse rate of the layer between 850 mb and 500 mb. In the west, VT has been shown to have a better correlation with thunderstorm activity. Although CT and VT can be used alone, when combined, they make the TTI a more reliable predictor of severe weather potential (AWS 1975). However, a forecaster must still use caution since a large TTI can occur due to a steep temperature lapse rate. The CT or low-level moisture should be evaluated to ensure there is enough moisture to support thunderstorm development. Table 2 demonstrates that the higher the value for TTI, CT, and VT, the greater the potential for severe thunderstorms.

2) K INDEX (KI)

The KI was developed by George (1960) to estimate the potential of air mass thunderstorms using the vertical temperature lapse rate combined with available moisture in the atmosphere. The KI is mainly utilized to predict heavy rain or indicate thunderstorm potential, not to make severe versus non-severe decisions (AFWA 1998). The following equation is used to calculate a value for KI:

$$KI = T(850\text{ mb}) - T(500\text{ mb}) + Td(850\text{ mb}) - [T(700\text{ mb}) - Td(700\text{ mb})]. \quad (4)$$

The standard units of the KI are Celsius degrees. The interpretation of the KI is the more positive the KI, the greater the potential for thunderstorm development. Table 3 displays a commonly used scale to determine thunderstorm occurrence. Notice that the KI applies to maritime tropical air masses (mT) and the Tropics.

3) SHOWALTER STABILITY INDEX (SHOW or SSI)

The Showalter Stability Index was originally developed by Showalter (1953) for the southwestern United States to determine thunderstorm potential based on the 850 mb to 500 mb layer. For the SSI, a parcel is lifted from 850 mb to the 500 mb level and the temperature of the lifted parcel is compared with the environmental temperature at 500 mb. If unsaturated, the parcel is lifted dry adiabatically until it reaches its lifted condensation level and from there it is lifted moist adiabatically to 500 mb. If the parcel is already saturated at the 850 mb level then it is just lifted moist adiabatically to the 500 mb level. The two values in Celsius degrees are compared to determine the stability of the layer. SSI values can be calculated using the following equation:

$$SSI = T_{environment}\ (500\ mb) - T_{parcel}\ (500\ mb). \tag{5}$$

Table 2. Range of values for thunderstorm prediction using TTI, CT, and VT (AFWA 1998).

Index	Region	Weak Thunderstorms	Moderate Thunderstorms	Strong Thunderstorms
Total Totals (TTI)	West of Rockies	48 to 51	52 to 54	> 54
	East of Rockies	44 to 45	46 to 48	> 48
	Europe	> 42	> 48	> 50
Cross Totals (CT)	East of Rockies	< 18 No TS	18 to 19	≥ 20
	Gulf Coast	< 16	20 to 21	
Vertical Totals (VT)	US: General			≥ 26
	Gulf Coast			≥ 23
	West of Rockies	< 28 No TS	28 to 32	> 32
	UK			> 22
	Western Europe			> 28

Table 3. Range of values for thunderstorm prediction using KI (AFWA 1998).

Index	Region	Weak Thunderstorms	Moderate Thunderstorms	Strong Thunderstorms
K Index (KI)	East of Rockies (mT / tropics)	20 to 26	26 to 35	> 35
	West of Rockies (mT)	15 to 21	21 to 30	> 30

The SSI is used as a first indicator of instability and performs best with well-developed storm systems in the Central United States (AFWA 1998). However, it is not a good index to use for severe weather when low level moisture is below 850 mb or when a frontal boundary or an inversion exists between the 850 mb and 500 mb layer.

The NWS commonly used the SSI as a measure of stability of the environment, not as a measure of severe weather potential. Table 4 shows the values of SSI used by the NWS. Notice as the values for SSI decrease, the potential for thunderstorms increases. SSI is one of the most common indices used in Air Force weather for determining the instability of the atmosphere (AFWA 1998). Table 5 shows the values commonly used by Air Force weather forecasters to interpret the potential for thunderstorms.

4) LIFTED INDEX (LI)

A problem was noticed with the Showalter index, so Galway (1956) developed the LI to solve for the potential weaknesses of the SSI. The weakness of the SSI is that it arbitrarily uses 850 mb as the starting point. The SSI is not representative of atmospheric stability in a region when there is an inversion present or a rapid drop in moisture between stations or two successive soundings. The LI uses the mean 3000-foot layer temperature and the LCL for the starting value to lift the parcel up to the 500 mb level. The following equation can be used to calculate a value for the LI:

$$LI = TEnvironment(500\text{ mb}) - TParcel(500\text{ mb}). \tag{6}$$

Table 4. Measure of stability using the SSI (NWS 2003).

Showalter Stability Index (SSI)	Stability
+1 to +2	Stable (Weak convection possible if strong lift present)
0 to -3	Moderately Unstable
-4 to -6	Very Unstable
< -6	Extremely Unstable

Table 5. Range of values for thunderstorm prediction using SSI (AFWA 1998).

Index	Region	Weak Thunderstorms	Moderate Thunderstorms	Strong Thunderstorms
Showalter Stability Index (SSI)	US	≥ +3	+2 to -2	≤ -3
	Europe	> 2 No TS	≤ 2 TS possible	

The temperature of the parcel lifted adiabatically to 500 mb is then compared with the environmental temperature at 500 mb. If the parcel 500 mb temperature is warmer than the 500 mb environmental temperature, the parcel will continue to accelerate upward, thus it is considered unstable. However, if the parcel 500 mb temperature is cooler than the 500 mb environmental temperature, then it will sink and be considered stable.

Table 6 shows the common values used by the National Weather Service for measuring the potential for convective activity. The LI values in Table 6 are based on the boundary layer temperature and moisture values. The degree of stability is determined from the values and then inferred to predict the potential for thunderstorms.

The Air Force version of the LI table applies the LI values directly to thunderstorm potential in Table 7.

5) SEVERE WEATHER THREAT INDEX (SWEAT)

The AWS and civilian community recognized the requirement for a stability index to both specify and predict areas of potentially severe convective weather (AWS, 1975). The SWEAT index was developed to integrate the thermodynamic and kinematic properties of the atmosphere into one index to indicate the potential for severe thunderstorms and tornadoes. Strict constraints were set on the new index to ensure the availability of a more timely and accurate forecast tool. The index must be:

1. Computed from upper air data,
2. Available one hour and 15 minutes after balloon launch,
3. Computed from current fields in data base,
4. Not dependent on derived parameters or complex pattern recognition.

Three versions of the SWEAT index were developed by Bidner (1970), Miller et al. (1971), and Miller (1972). The original version developed by Bidner did not include the shear term that was added to the second version, which is most widely known and used. The third version of SWEAT relocated the height at which the dew point temperatures were measured, from 850 mb to 900 meters. The shear measurement was changed from the 850 mb level to a step function to take into account shear of the layer

Table 6. Common values of Lifted Index (LI) used for predicting atmospheric stability (NWS 2003).

Lifted Index (LI)	Stability
LI over 0	Stable but weak convection possible for LI= 1-3 with strong lift
LI = 0 to -3	Marginally unstable
LI -3 to -6	Moderately unstable
LI -6 to -9	Very unstable
LI below -9	Extremely unstable

Table 7. Range of values for thunderstorm prediction using LI (AFWA 1998).

Index	Region	Weak Thunderstorms	Moderate Thunderstorms	Strong Thunderstorms
Lifted Index (LI)	All	0 to -2	-3 to -5	-5 and lower

from 900 meters to 500 mb. For this research, the focus will be on the second, most widely used version of SWEAT, which is given by the following formula:

$$SWEAT = 12(A) + (20)(TTI - 49) + 2(f8) + f5 + (125)(S + 0.2), \quad (7)$$

where A is the dew point at 850 mb, TTI is the result found from the Total Totals Index, f8 is the 850 mb wind speed (in knots), f5 is the 500 mb wind speed (in knots), and the shear term S is the sine of the angle of the 500 mb wind direction minus the 850 mb wind direction. If a term in the equation is negative, the term is set to zero. It is also important to know that the shear term is set to zero if the following criteria are not met:

1. 850 mb winds are between 130 to 250 degrees.

2. 500 mb winds are between 210 to 310 degrees.

3. 500 mb wind direction minus the 850 mb wind direction is not positive.

4. The 850 mb and 500 mb wind speeds are at least 15 knots.

The SWEAT index was created to distinguish between severe and non-severe storms, so it should not be used to forecast for general thunderstorms (AWS, 1975). However, the SWEAT index was used in this research since the focus is on the prediction of CG lightning from all thunderstorms.

The Air Weather Service performed a case study to test the performance of the SWEAT index on past occurrences of severe weather. The past cases were grouped into two categories: tornadoes or other severe weather criteria. For the study, 159 severe storms were used and the values in Table 8 are the results of the study. See AWS (1975), for more detailed information on the study performed. False alarm rates were not taken into account since the study only included storms that contained severe criteria.

6) KO INDEX (KO)

The German Weather Bureau developed the KO index in an effort to better predict thunderstorms in the European region. The KO index is responsive to the amount of moisture present and performs best in the cooler climates of Europe and the Pacific Northwest. The complexity and method for calculating the KO index are the major drawbacks to its use. The following equation is used to calculate the KO index.

$$\mathrm{KO} = \frac{\theta_e(500mb) + \theta_e(700mb)}{2} - \frac{\theta_e(850mb) + \theta_e(1000mb)}{2} . \tag{8}$$

where θ_e is the equivalent potential temperature for each level. θ_e can be found by locating the LCL for the level of interest and continuing up moist adiabatically until all moisture is removed, where the moist and dry adiabats are parallel. Then continue from there dry adiabatically to the top of the chart and read the value (AFWA, 1998). Table 9 was developed for USAF weather forecasters to apply the values of the KO index towards thunderstorm forecasting.

7) CONVECTIVE AVAILABLE POTENTIAL ENERGY (CAPE).

CAPE involves the vertical integration of the atmospheric profile to determine a measure of the maximum possible energy available for buoyancy and updrafts. CAPE is measured from the level of free convection (LFC) to the equilibrium level (EL) on the Skew-T log P diagram. The parcel can be lifted from the boundary layer or the surface, therefore, the value of CAPE depends on the height from which the parcel was lifted. The higher the value of CAPE measured, the greater the potential for thunderstorms.

Table 8. The accepted values of the SWEAT index for severe weather (NWS, 2003).

Severe Weather Threat (SWEAT) Index	**Thunderstorm Potential**
SWEAT over 300	Potential for severe thunderstorms
SWEAT over 400	Potential for tornadoes

Table 9. Range of values for thunderstorm prediction using KO (AFWA, 1998).

Index	Region	Weak Thunderstorms	Moderate Thunderstorms	Strong Thunderstorms
KO-Index (KO)	Cool, moist climates: Europe, Pacific NW	> 6	2 to 6	< 2

The shape of the positive area is a critical detail about CAPE that must be considered when trying to determine the type of thunderstorm that could potentially develop. A long, thin positive area represents a slow vertical acceleration but higher tops where a short, fat positive area involves rapid acceleration. A tall, thin profile suggests the potential for high precipitation thunderstorms and a short, fat profile indicates a potential for rotating updraft development (NWS, 2003). Table 10 shows the values considered by Air Force Weather forecasters when determining the potential for thunderstorms (AFWA, 1998). As with other parameters, the NWS uses slightly different values than AFWA and uses the values as a stability indicator for its forecasting purposes. Table 11 indicates the values a NWS forecaster would consider when using CAPE as an indication for thunderstorm potential. The following equation is one of the variations used to determine the values of CAPE:

$$CAPE = g\int_{LFC}^{EL} \frac{(T_{Parcel} - T_{Environment})}{T_{Environment}} dz \quad . \tag{9}$$

where g is gravity and T represents the temperature for each level. The units for CAPE are Joules per kilogram. Another variation of CAPE uses virtual temperature instead of temperature.

CAPE is technically not a stability index like the SI and LI since the units of CAPE are energy and not temperature and the standard indices compare temperature differences between layers (Blanchard, 1998). In his study, Blanchard (1998) found only moderate correlations between other common stability indices and CAPE. He concluded the low correlations were a function of the difference between what the indices were actually measuring in the atmosphere. However; the NWS (2003) often considers CAPE to be better than standard indices for indicating instability since CAPE is calculated by integrating the whole layer and is not dependent upon a certain level. CAPE is included in this research as a possible predictor of CG lightning since it is used throughout the weather community as a tool to determine instability and the possibility of thunderstorms.

Table 10. Range of values for thunderstorm prediction using CAPE (AFWA, 1998).

Index	Region	Weak Thunderstorms	Moderate Thunderstorms	Strong Thunderstorms
CAPE (J/kg)	East of Rockies	300 to 1000	1000 to 2500	2500 to 5300

Table 11. The accepted values of CAPE (J/kg) for determining stability (NWS, 2003).

Convective Available Potential Energy (CAPE)	Stability
CAPE below 0	Stable
CAPE = 0 to 1000	Marginally unstable
CAPE = 1000 to 2500	Moderately unstable
CAPE = 2500 to 3500	Very unstable
CAPE above 3500-4000	Extremely unstable

3. Data and Methodology

a. Overview

One of the goals for this research was to find a user friendly prediction tool for CG lightning for each of the locations in 15th OWS AOR. Keeping this main objective in mind, the predictors were limited to a group that could easily be created and used by the forecaster in the field. All the data used in this research were provided by Air Force Combat Climatology Center (AFCCC). The data set consisted of CG lightning data from the NLDN, surface data provided from each location, and upper air data obtained from the rawinsonde sounding data. A major concern not previously considered by Venzke (2001) was the location of fronts, troughs, or any other synoptic weather feature in relation to the upper air station. By separating the point of interest from the upper air station, the variation in weather across a distance as much as 200 miles could decrease the value of the results achieved in this research.

Sahu and Singh (1999) tried using over 128 various single station predictors, with a maximum model of 12, to forecast the categorical occurrence of thunderstorms in India. They used two types of data for the research, hourly surface observations and Radiosonde data for 00Z and 12Z, and considered the forecast area as a 100 km radius. The regression equations were validated with independent data, and the results provided a probability of detection of 72.5% for 00Z and 70.9% for 12Z, and a false alarm rate of 39.3% and 37.1% respectfully. Relying on surface observation for thunderstorm verification was one limitation to the research. Some thunderstorm activity was missed due to the spacing of the locations combined with the mesoscale nature of thunderstorms. The results achieved were promising considering the limitations. This research took a similar approach by using both surface and upper air data but used the NLDN to obtain a more accurate count of CG lightning to use as the predictand. The knowledge obtained from both Venzke (2001) and Sahu and Singh (1999) helped in the selection process for the predictors used in the decision tree building process. The following rules were required to ensure the forecast tool would be valuable to forecasters in the field: 1) the predictor was created from data that would be readily available to forecasters in the field and 2) the calculations were required to remain relatively simple to ensure the timeliness of the forecast tool.

Reap (1994) also did research on predicting thunderstorms using prediction tools. He developed experimental thunderstorm probability equations based on large-scale synoptic forcing. He used the K stability index combined with lightning frequencies as predictors for the research in which he used pattern classification to develop map patterns to create climatology of lightning distributions for each synoptic situation. In the study he focused on the "warm season" period of March thru September. The following predictors were used in the study by Reap (1994) and in this research: wind, surface pressure, Total Totals Index (TTI), K index (KI), wind shear, and temperature advection.

b. Data

Choosing the best data set for this research involved determining the best time of year so the variability of the upper air stability indices is great enough and there were enough CG lightning strikes to draw reasonable conclusions. The 15^{th} OWS AOR followed a seasonal pattern with June and July as the peak months. Thunderstorm damages in the northern regions of the U.S. also had a narrow distribution centered on the summer months (NWS 1997). Reap (1994) showed that even for Florida a lightning study can focus on the warm season due to the normal distribution of the CG flash count per day centered over June and July. Venzke (2001) also concluded that the frequency of lightning is significantly lower during the cool months October thru April, therefore a warm season (1 March thru 30 September) data set was chosen. Table 12 shows the locations matched with the upper air stations used for this research.

1) SURFACE DATA

Venzke (2001) focused on stability indices in his research, but as mentioned before, his research was based on one location, not separate locations as in this research. Therefore, surface data were included to add more predictors to help offset the shortcomings of using just upper air data. The surface winds and station pressures at the points of interest were used in comparison with the same data at the upper air station to

Table 12. 15th OWS locations matched with the upper air stations used in this research.

Location	UPPER AIR ST NAME	LAT	LON	ELEV
Dover	724030 VA WASHINGTON/DULLES	38.983	-77.467	95 m
Andrews	724030 VA WASHINGTON/DULLES	38.983	-77.467	95 m
Ft Belvoir	724030 VA WASHINGTON/DULLES	38.983	-77.467	95 m
McGuire	725200 PITSBURGH	40.52	-80.22	359m
WPAFB	724260 OH WILMINGTON RAOB	39.417	-83.817	323 m
WPAFB	724290 OH DAYTON/JAMES M COX	39.867	-84.117	298 m
Westover	725180 NY ALBANY COUNTY ARPT	42.75	-73.8	86 m
Ft Drum	725280 NY BUFFALO INTL ARPT	42.933	-78.733	218 m
Scott	725320 IL PEORIA REGIONAL	40.667	-89.683	201 m
Scott	745600 IL LINCOLN UPPER-AIR	40.15	-89.333	178 m
Offutt	725530 NE NORTH OMAHA	41.367	-96.017	399 m
Offutt	725580 NE OMAHA/VALLEY	41.317	-96.367	350 m
Ellsworth	726620 SD RAPID CTY RGNL ARPT	44.067	-103.217	1029 m
Grand Forks	727640 ND BISMARCK MUNICIPAL	46.767	-100.75	505 m
Minot	727680 MT GLASGOW INTL ARPT	48.217	-106.617	696 m
Grissom	744550 IA DAVENPORT UPPER-AIR	41.617	-90.583	229 m

determine similarities in air mass. The difference in station pressure from the upper air site and the surface station and the three hour surface station pressure change was also added to the problem to determine if they could be used as a CG lightning prediction tool. Using the extra surface terms provided the research with several more predictors, which could possibly increase the value of the findings. Determining frontal location manually was considered. However, frontal location is very subjective and locating daily surface charts for 11 years proved an overly difficult task.

2) UPPER AIR DATA / STABILITY INDICES

The stability indices used in this research were computed by AFCCC using archived upper air data and FORTRAN algorithms written specifically to determine the value for each index. Various quality control measures were implemented to limit the number of errors in the data set. The data set was manually checked for errors. A value threshold formula was applied to each index to verify that the indices were in the acceptable range. For example, negative values were not accepted for TTI and CAPE.

The upper air data were obtained from rawinsonde balloon flights from the stations listed in table 13. The rawinsonde measures the atmospheric column and records the following data: temperature, pressure, dew point, wind speed, and wind direction. The measurements are assumed to be directly overhead from the launch location, however the balloon actually drifts with the atmospheric winds as it ascends. There are two main sources for error in the rawinsonde data: 1) the balloon does not rise at the standard rate of 300 meters per minute, and 2) a lag time in the sensors (Allen 2003). The pressure and temperature lag time are of major concern, since both pressure and temperature are the primary items used to determine stability indices. Golden et al. (1986) determined that the lag time for temperature sensors was 4 to 20 seconds and increased with height. Golden et al. (1986) also concluded that the standard error for pressure was ± 1 mb at the surface and ± 2 mb at 500 mb.

3) LIGHTNING DATA

The lightning data used in this research were cloud-to-ground lightning strikes as reported by the NLDN to be within 25nm and 10nm of the 13 locations previously defined under the watch of the 15th OWS. The main focus of this research was the forecasting of CG lightning. There were no attempts to determine the severity of the individual storms since it is beyond the scope of this research. Also, due to limitations in the sensors used in the detection of lightning, no intra-cloud lightning can be inferred in this study (Cummins et al., 1998). Orville and Huffines (1998, 2000) graphically displayed normal curves for flash counts which centered over June and July. They also showed that the months of January, February, October, November, and December contained very low flash counts. So, to ensure availability of enough CG lightning data, this research will focus on the months March thru September. Wacker and Orville, (1999a) concluded the data before and after the 1994 upgrade did not contain significant ambiguities and are acceptable for research.

Venzke (2001) concluded that there were too few lightning strikes within 10 and 25 nm to develop a relationship, so he used lightning data within 50 nm to represent the 13 upper air stations chosen in his research. The 15th OWS was more concerned with lightning within a smaller radius so this research focused on 25 nm and 10 nm.

The CG lightning data were grouped into two categories to match the upper air data. The 00Z lightning count consisted of CG lightning that occurred within a 25 nm and a 10 nm radius of the location from 0000Z to 1159Z and the 12Z lightning count contained CG lightning within a 25 nm and a 10 nm radius of the location from 1200Z to 2359Z.

c. Statistical Methodology

Statistics evaluates and quantifies uncertainty and makes inferences and forecasts about the uncertainty (Wilks 1995). Simple linear regression is the comparison of two variables to fit a linear equation to the observed data. The data should be compared using a scatterplot to see if there is a relationship between the variables of interest. If there is no noticeable correlation, then a linear regression model would not provide a valuable tool.

Since weather uses a number of stability indices for thunderstorm prediction, a different method was needed to account for any correlation between multiple variables. The method of multiple linear regression was applied to the data sets to determine if a usable tool could be found. Multiple linear regression is a method of regression analysis that involves more than one regressor variable (Montgomery and Runger 2003).
Linear regression was not chosen for this research because: 1) predictors and predictands forecasting weather outcomes typically do not display a linear correlation, 2) linear regression equations are difficult to use in weather forecasting due to the number of variables, and 3) multiple outcomes are possible for the same value of the predictor. The

number of variables chosen to predict CG lightning in this research created a very complicated multivariable regression equation. The goal of the research was to find a technique to predict CG lightning to be used in the forecast community, therefore the results needed to be user friendly and timely. So for this research, a newer statistical tool was used to produce a valuable prediction model. However, the results from the new model will be validated using classical statistics (see chapter 4 for results).

d. CART

1) CART OVERVIEW

Classification and regression tree (CART) analysis, developed by Salford Systems, utilizes the tree building technique to define the complex relationships between several predictors and the target values (predictands). A major break through in statistical evaluations, CART is a computer intensive technique that can automatically analyze any number of variables, regardless of missing data regions or outliers, and create a simple yet effective model. CART analysis takes advantage of a binary recursive partitioning system which divides a major category known as "nodes" into two child nodes.

CART uses a decision tree to represent the entire model or analysis as determined from the interaction between the relevant predictors and the predictands. The main nodes divide to child nodes using rules established by the model to determine the proper split at each level. The questions used to determine the split are formulated so that only two answers are possible. For example is SSI less than 0? If yes, the case will split to the left or if no, the case will take the right hand path. The path ends at a terminal node at which time the object is classified and all the rules used to reach the terminal node can be reviewed. This method supplies the end user the rules to follow from the initial to final decision for solving their problem. CART provides a simpler interpretation than a multivariate regression model making it an ideal tool to use in the weather forecast decision process.

(i) Advantages of Using CART as a Statistical Tool

CART is non-parametric because no assumptions are made about the distributions of the data sets. Therefore it can work with numerical data that are highly skewed or multi-modal and also work with categorical predictors as well. Researchers save time that otherwise would be used to determine the normality of the data and to transform data sets that are not normal. CART searches all possible variables to determine the splits required to make the decisions even if the problem contains hundreds of variables. CART can create decision trees even if some of the predictor variables are missing from certain areas of the data. "Surrogate" variables which contain similar information to the actual variables are used in place of the missing variables. Thus decision trees made for data sets with missing data use both the actual and the surrogate variables to determine the splitting. CART analysis is a relatively automatic method which requires very little input from the user, unlike multivariate models which require extensive input, analysis of early results and modification of the methods used to create a useable model. CART results are displayed as trees which the end users understand easier since the structure of the rules and their logic is apparent.

(ii) Disadvantages of Using CART as a Statistical Tool

CART analysis is new and not well-known, therefore some traditional statisticians are reluctant to accept the process. Due to earlier poor performances and unrealistic claims by other tree methods, CART must overcome a common distrust. CART help is difficult to find since it is not a standard technique and it is not included in many major statistical software packages.

2) CART CLASSIFICATION TREES

Lewis (2000) discussed the four components to a classification problem: 1) a categorical outcome or dependent variable which one wishes to determine, 2) the predictor variables or independent variables and the possible relationship with the outcome, 3) the data set which contains the values for the outcome and each of the predictor values, and 4) the test data set for validating the decision rules created. CART takes a categorical data set and develops a decision tree based on all possible combinations of the data. The CART program allows for user input during certain phases in the tree building process. The following were used in this research to help develop the most reliable decision tree based on the information provide by the user manual.

(i) Cost Matrix

A cost matrix is one way CART allows the user to adjust the importance of the categorical predictands. To determine the cost matrix for the CG lightning problem, one must decide if it is equally as bad to misclassify thunderstorm occurrence versus non-occurrence. To weigh the costs equally will reduce the false alarm rate, however it will also reduce the predictability rate for CG lightning. For this research two cost matrix methods will be applied to each location; one method that optimized the prediction of CG lightning but at the expense of higher false alarm rates and the other method that used equal costs to give the maximum predictability rate while also focusing on reducing false alarm rates.

(ii) Variable Importance

An advantageous benefit of using the CART analysis program to data mine is the ability of CART to determine strictly from the data a numerical grouping of the predictors for each decision tree based on their importance to the outcome. For each decision tree created, the user can easily determine which variables are important and which ones are not. From this information, the user can eliminated the useless predictors and rebuild the decision tree with only the predictors determined by CART to be of prediction value.

(iii) Priors

The priors tab of the CART program allows the user to choose the influence of the data set on the overall model. There are six user choices 1) priors equal, 2) priors learn, 3) priors test, 4) priors data, 5) priors mix, and 6) priors specify. The choice of priors is based on the data set and user requirements. Priors equal make the probability of each class occurring equal. With priors learn CART matches the probabilities with the learn sample frequencies. Similarly, for priors test, CART matches the probabilities with the test sample frequencies. Priors data uses the probabilities of the total sample frequencies. Priors mixed combines the average of the priors equal and priors data. Finally, priors specify allows the user to set the probabilities of each class.

This research used priors equal since this selection gives each class an equal chance to be classified correctly. The other methods were discounted for various reasons. The CART user manual described priors data as least likely to give positive results when focusing on a class with lower frequency (Salford Systems 1995). Priors data will cause CART to focus on correctly classifying the larger class. For this research, CG lightning frequency average approximately 10%, therefore priors data or any combination of it was eliminated as an option since it would focus on the nonoccurrence of CG lightning. This decision eliminated all choices excepted priors equal and priors specify. To eliminate human error, priors specify was also not chosen since it required the user to make a decision that would affect the entire tree building process.

3) SPLITTING RULES OF CART

CART allows the user to choose from six different splitting methods: Gini, Symmetric Gini, Entropy, Class Probability, Twoing, and Ordered Towing. The user must decide which rule is best for the data based on knowledge of each tool. The following explanation on splitting rules can be found in the CART manual (Salford Systems 2002).

(i) Gini

The Gini method is the standard splitting rule that splits the data to obtain a one class prevailing terminal node. The Gini method is considered the best for a two-level dependent variable that can be predicted with a relative error less than 0.5. Costs are incorporated in the Gini method by adjusting the prior probabilities before the tree is grown.

(ii) Symmetric Gini

The Symmetric Gini method differs from the Gini method only in the way the method handles cost. The costs in the Symmetric Gini method are made symmetric and incorporated into the impurity function at the tree growing stage.

(iii) Entropy

The Entropy method looks for splits where as many levels as possible are divided as perfectly as possible. As a result, Entropy emphasizes getting rare events correct relative to the common events.

(iv) Class Probability

The Class Probability forces CART to build probability trees instead of classification trees. A class probability tree attempts to separate segments of the data by different probabilities of response.

(v) Twoing

The Twoing method differs from the Gini method on the multilevel targets. Twoing generates more even splits separating whole groups of classes. Twoing tends to be best for target variables with four to nine levels and for two-level dependent variables predicted with a relative error of 0.8 or better.

(vi) Ordered Twoing

Ordered Towing is a variation of the Twoing method designed for ordered targets. The rule will group target classes that are adjacent to each other. For example, (1,2,3,4,5,6) would be grouped (1,2,3) and (4,5,6). Ordered Twoing works best with numeric levels and does not work well when the target is a character.

The six splitting rules were applied to the CG lightning data sets to determine which method would provide the most beneficial outcome. The CART user's manual provided several rules of thumb which were considered as the testing phase was performed. Based on the size and type of data sets used and the results desired for this research, the rules of thumb supplied in the CART user's manual provided guidance to eliminate three of the splitting rules: Symmetric Gini, Class Probability, and Ordered Twoing. For most of the data sets, there was very little difference between the Gini, Twoing, and Entropy methods. However, overall the Gini method produced the best results when a difference could be determined and will be used as the primary method for CART analysis.

4) DATA VALIDATION

The categorical trees were validated using the cross validation technique and an independent data set containing 2003 data. Freestrom (2001) stated the cross validation method had several advantages over other methods. 1) The cross validation method does not hold out data. When the data set only contains a few hundred occurrences of CG lightning, holding out random data would greatly reduce the effectiveness of the model.

2) The user can choose the number of cross validation trees desired to perform the tests. The cross validation method builds the chosen number of trees and compares the optimal tree with the other trees built.

CART also gives the user the opportunity to test the trees using an independent data set. The test data can be used during the tree building process or they can be used to test the tree after the building process is complete. Both cross validation and independent data validation methods were used. Since cross validation uses a random portion of all the data to validate the trees, independent data were chosen to test the decision trees to represent a year of true verification for the forecast tool. This decision tool will be placed in the field with no yearly updates to the trees, so using an independent data set allowed for an operational test without the costs involved with a missed forecast. Using 2003 data for validation provided a better insight to the feasibility of decision trees in the field without risking life or equipment.

4. Results

a. Introduction

This chapter presents a summary of the decision tools created by the CART program to forecast CG lightning. Four CART decision trees were produced for each of the 13 locations. The decision trees were developed using two different radii and cost matrixes. Two data sets were used in this research. The learning data set is from 1993 – 2002 and was used to build the models. The test data set is from 2003 and was an independent data set used to validate the models. As previously mentioned, this research was verified using the independent data to "field test" the forecast decision trees. So it is important to discuss the results from both decision trees since 2003 could be an extreme year. A tree that verifies poorly but has a good learning score could validate well in 2004. The results will demonstrate the varying success of the decision tools for each location and the benefits to the weather forecaster.

b. Variable Importance

Variable Importance was used to maximize the decision tool while reducing the number of predictors required to achieve the best results. As mentioned in chapter 3, variable importance allows the user to quickly determine the predictors that supply value to the decision tree. CART displays the predictors in order of relevance in the summary reports. Variable importance also supplied the research with insight on which variables worked best with each region. Table 13 contains the top two predictor variables of each tree created for each location forecasted. Based on the results, SSI, KI, and LI were the top predictors for the research.

c. Decision Trees

The decision trees built in this research took advantage of CART allowing the user to input the importance of each class. The cost matrix is the tool that allows the user to decide the importance of correctly forecasting CG lightning. A one to one cost weighed the importance of correctly forecasting CG lightning and false alarm rates equally. A tree built using a two to one cost matrix expressed the importance of correctly forecasting CG lightning. These trees simply implied that it was twice as important to not miss a CG lightning occurrence as it was to have incorrectly forecast for CG lightning and it not occur. Fig. 3 is an example of a decision tree created by CART. The tree was for within a 25 nm radius of Andrews AFB using a cost matrix of two to one. The SSI was the only predictor used as a rule and the split was performed at 3.66. The SSI was slightly higher than the "accepted" values of SSI in chapter 2. The higher value of SSI could account for the over forecasting observed in the verification tests.

Applying a cost matrix to the data analysis process did improve the prediction success for most locations. However, as stated above and in chapter 3, the higher successes for predicting CG lightning increased the misclassification error for the

Table 13. Top two predictors for each location based on the variable importance. The following indices are listed below; Showalter Stability Index (SSI), K Index (KI), Lifted Index (LI), Convective Available Potential Energy (CAPE), Month, and Severe Weather Threat Index (SWEAT).

Location	**10 nm 1 to 1 cost**	**10 nm 2 to 1 cost**	**25 nm 1 to 1 cost**	**25 nm 2 to 1 cost**
Andrews AFB	SSI KI	SSI KI	SSI KI	SSI LI
Dover AFB	LI SSI	SSI LI	SSI LI	SSI LI
Ellsworth AFB	LI SSI	LI SSI	SSI LI	SSI LI
Ft Belvoir	SSI KI	SSI KI	SSI KI	SSI KI
Ft Drum	LI CAPE	LI SSI	LI SSI	LI CAPE
Grand Forks AFB	SSI LI	SSI LI	LI SSI	LI SSI
Grissom AFB	KI SSI	KI SSI	KI SSI	SSI KI
McGuire AFB	CAPE SSI	SSI LI	KI SSI	SWEAT SSI
Minot AFB	LI SSI	LI Month	KI LI	KI LI
Offutt AFB	KI SSI	KI SSI	KI SSI	KI SSI
Scott AFB	KI SSI	KI SSI	KI SSI	KI SSI
Westover AFB	KI SSI	SSI KI	KI SSI	KI SSI
Wright Patterson AFB	SSI LI	SSI KI	SSI KI	SSI KI

non-occurrence of CG lightning. So the increased prediction capabilities came with the price of increased false alarm rates.

A problem with the CART decision tree is determining which node holds the most important information and is the best tool for forecasting the desired outcome. There are four probabilities used to determine the value of each terminal node in this research: 1) total probability of the learning tree, 2) total probability of the test tree, 3) the within the terminal node probability of the learning tree, and 4) the within the terminal node probability of the test tree. The "total probability" (TP) is calculated by adding the probability of each class for the specific node.

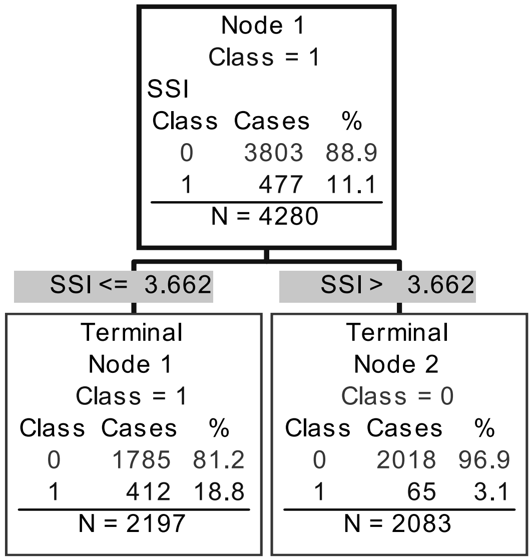

FIG 2. A learning decision tree for within 25 nm radius of Andrews AFB with a two to one cost matrix.

The classes are: (class 1) CG lightning occurred and (class 0) no CG lightning occurred. The probability within each terminal node is calculated using Bayes Rule. This probability displays how accurate the prediction is within each terminal node and can be used to measure the strength of the forecast. The following equation is used calculate total probability (TP):

$$TP = P(left|class1)P(class1) + P(left|class0)P(class0)\,. \tag{10}$$

The following equation is Bayes Rule (Montgomery and Runger 2003):

$$P(class1|left) = \frac{P(left|class1)P(class1)}{P(left|class1)P(class1) + P(left|class0)P(class0)}\,. \tag{11}$$

The direction can be left or right depending on the splitting rules used. For both equations, left is the direction of the branch in the decision tree, class 1 is CG lightning occurred, and class 0 is CG lightning did not occur. Next, an example calculation of both probabilities using the following decision tree was performed to verify the CART numerical output.

First calculate the individual portions:

$P(SSI <= 3.662 | class1) = \frac{412}{477}$, $P(SSI <= 3.662 | class0) = \frac{1785}{3803}$, and the probability of each class is 0.5.

Now using equations 10 and 11:

$$TP = (0.8637)(0.5) + (0.4694)(0.5) = 0.6665 \text{, and}$$

$$P(class1 | SSI <= 3.662) = \frac{(0.8637)(0.5)}{(0.8637)(0.5) + (0.4694)(0.5)} = 0.64791.$$

The total probability for terminal node one is 0.6665 and the within node probability is 0.64791. These results imply that there is a 66.65% probability of finishing in terminal node one. Once in terminal node one, there is a 64.791% chance of properly forecasting CG lightning. This is the exact result that CART presented in the output. The TP results were used to reduce the decision trees to contain only the relevant rules. If the TP was less than 0.01, meaning there was a less than 1% probability of the node occurring, then the node and it's accompanying rules were removed. Appendix B contains the forecast rules and their calculated probabilities using Bayes' rule. Displaying the probabilities provided the reader the probability of correctly forecasting CG lightning for each forecast rule.

Classical statistical methods were not used to create the forecast tools; however, classical statistics will be used to determine the value of the final decision trees. Contingency tables were created to evaluate the categorical forecast results. A contingency table is a simple two by two that displays the forecasts and the observations in an easy to understand method (Wilks 1995). Six verification tests were applied to the resulting binary forecasts located in the contingency tables. Fig. 2 is a base table with the box identifiers for applying to the formulas from the contingency tables. All the values for the six verification equations can be found using Fig. 2.

	Observed	
Forecasted	YES	NO
YES	A	B
NO	C	D

FIG 3. A contingency table showing the correct location for the results of binary forecasts (modified from Wilks 1995).

The hit rate was the first verification method applied to the results. Hit rate is the number of events or nonevents forecasted correctly divided by the total number of forecasts (Wilks 1995). The best hit rate is a score of one and the worst is a score of zero.

The hit rate (H) was calculated using the following equation:

$$H = \frac{A + D}{n}, \qquad 10$$

where n is the total forecasts given by adding together all four boxes in the table.

The second method applied to the results was a threat score (TS). The TS is a technique advantageous to determine when the number of positive or yes events occur less frequently than the nonevents (Wilks 1995). Since CG lightning occurrence was one-tenth of the nonoccurrence this test was also calculated from the results. Again similar to the hit rate, a score of zero is the worst and a score of one is the best possible score. The TS was computed using the following equation:

$$TS = \frac{A}{A + B + C}. \qquad 11$$

The third verification tool applied was the probability of detection (POD). POD is probability the event would be forecasted given that the event actually occurred (Wilks 1995). Perfect forecasts have a POD of one and the worst forecasts have a POD of zero. The following equation was used to determine the POD:

$$POD = \frac{A}{A+C}. \qquad 12$$

The previous tests were used to judge how good the forecasts verified; however, the fourth method determined how many forecasts did not occur. The false-alarm rate (FAR) was calculated for the decision trees to clarify if the high forecast results were obtained by extreme over-forecasting. Wilks (1995) presented the following formula for evaluating FAR:

$$FAR = \frac{B}{A+B}. \qquad 13$$

A major difference of FAR over the previous methods is smaller values are better with zero as the best score and one the worst score.

The Bias was the fifth method used to verify the results. Bias compares the average of the observations and forecasts to determine a ratio (Wilks 1995). The resulting ratio allows conclusions to be made about the forecast model. A perfect model has a Bias of one. If a model has a Bias greater than one, it is said to over-forecast the event since the model forecast the event more than the event occurred. A Bias less than one describes a model that under-forecasts the events. Bias was calculated using the following equation:

$$Bias = \frac{A+B}{A+C}. \qquad 14$$

A Kuipers Skill Score (KSS) was the last measure used to determine the value of the forecast model. A skill score is a scalar measure of how well the forecast model performed (Wilks 1995). Most weather forecast models are compared with some type of climatology to determine the benefits of choosing the particular model. So, the KSS was chosen for this research because the random reference forecasts in the KSS formula contain a marginal distribution equal to climatology (Wilks 1995). The equation used to determine the KKS was:

$$KSS = \frac{AD - BC}{(A+C)(B+D)}. \qquad 15$$

A score of one is a perfect forecast, a score of zero is a random forecast, and a negative score is a forecast that is inferior to a random forecast.

Table 14 is a complete contingency table with verification of the model results for within 25 nm of Andrews AFB and a cost matrix of two to one (using the forecast decision tree in Fig. 3). The data used to build this model were for 1993 – 2002. Table 15 is the same as table 14 except the tests were performed on the verification data from 2003.

The verification tests were performed on the both decision trees to determine their value as a forecast tool. There were 822 cases of CG lightning during the 10year period. The decision tree correctly classified 736 of those occurrences. This gave the tree a POD of 89%, however, the model over forecasts which can be seen the high FAR of 0.68. The over forecast was apparent with a Bias of 2.8.

Table 14. Model results for Andrews AFB 25 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	736	1598	61%	0.30	0.89	0.68	2.8	0.43
NO	86	1860						

Table 15. Verification results for Andrews AFB 25 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	109	142	66%	0.43	0.98	0.57	2.3	0.53
NO	2	175						

The first five verification tests applied to the results gave mixed results so a skill score technique was calculated to determine the usefulness of the decision tree. A KSS of 0.43 showed the tree his was an improvement over a random forecast. The verification tree, based on 2003 data, performed better as seen by the numbers in Table 15. There were 111 Cg lightning occurrences. The decision tree when applied to 2003 had a POD of 98% and a FAR of 0.57. The bias of 2.3 was lower than the 2.8 of the base tree. A KSS of 0.53 demonstrated that the model did really well when compared to random forecasts. Complete contingency tables with results from the six verification tests can be found in Appendix A.

5. Conclusions and Recommendations

a. Conclusions

Thunderstorms are weather phenomena that affect all of North America. A major product of these thunderstorms, cloud-to-ground (CG) lightning is extremely important to the weather and Air Force communities. The world's increasing dependence on technology and energy has increased the importance for the weather community to achieve the ability to predict the occurrence of thunderstorms and CG lightning. The Department of Defense (DoD) and the Federal Aviation Administration (FAA) are extremely concerned with CG lightning due to the risk to personnel and the cost of replacing damaged equipment. With a single aircraft costing millions of dollars, the need for better thunderstorm and lightning warnings is becoming more crucial than ever.

The 15th Operational Weather Squadron (OWS) located at Scott AFB, Illinois has the responsibility of forecasting weather 12 active duty and 17 Reserve and National Guard military installations across the North Central and Northeastern United States. The 15th OWS made a request for a reliable forecast tool to predict CG lightning. To fulfill their request, forecast decision tools were created using a Classification and Regression Tree (CART) data analysis program.

Venzke (2001) developed a similar method of applying stability indices computed from upper air soundings to determine the probability of CG lightning for a region within a 50 nautical-mile (nm) radius surrounding the upper air stations in the Midwest. However, Venzke did not apply the CART data analysis program to his work. Also, these locations were geographically separated from the upper air stations, in some cases by over 100 miles. The primary purpose of this research was to provide the 15th OWS with a CG lightning forecast tool using predictors that were readily available to the forecasters. Forecasters at most military installations do not have the benefit of on-site upper air stations so, this research used the resources available to the military operational forecasters in an effort to determine a predictor for CG lightning for within 25 nm and within 10 nm of each location. Knowledge learned from both Venzke (2001) and Sahu and Singh (1999) was implemented into the process to determine the best possible methods for developing the CG lightning forecast decision tools.

The occurrence of CG lightning for the North Central and Northeast regions is centered on the summer months, therefore this study focused on a data set containing the warmer months. A 10 year data set from 1993 to 2002, using only the months March through September was determined to represent the region of interest. To validate the decision model, an independent year was chosen as the verification data. 2003 data were applied to the model to establish the validity of using the model as a forecast tool in the field. The CG lightning data were collected from the NLDN. Two criteria were used, within 25 nm and within 10 nm of each location. The upper air sites were selected to best represent the conditions for each location. The stability indices were calculated and combined with the lightning data. Later surface data was added to create a larger list of predictors.

CART was the tool used in this research to create forecast decision aides for the locations within the 15th AOR. The CART data analysis program developed decision rules that were easy to follow and can be implemented into an automated system. CART was applied to a modified process similar to the one performed by Venzke (2001) and ideas for new predictors were taken from the works of Sahu and Singh (1999). CART allowed for many user interactions into the forecast model. Two cost matrices were used to emphasize the importance of correctly forecast CG lightning. Priors data was chosen to assign the probabilities to the data so that the occurrence of CG lightning was as probable as the non-occurrence. This prevented CART from focusing on the non-occurrence. The Gini splitting method was chosen to make the splits in the decision trees. Independent testing was selected over cross validation although both methods were used to validate the models. The independent data allowed for a trial field test without the impact of missed forecasts. The verification using 2003 data provided insight on the true value of the CART decision trees as true forecast tools.

Four decision trees were produced for each location using the period of record from March through September, 1993 to 2002. CART compared the upper air stability indices and surface data at 12-hour intervals with CG lightning data occurring within the next 12 hours to determine prediction rules. Once the decision trees were created, the data from 2003 were used to independently verify of the decision trees. The variable importance was applied to each decision tree created and the final 52 trees were accepted and tested for using statistical verification methods. The top predictors were SSI, LI and KI. They were listed as the top two predictors for nearly every decision tree built (see table 13). A surprise predictor variable surfaced in certain regions of the 15th. CART used the month as a predictor for Ellsworth AFB, Grand Forks AFB, Grissom AFB, Minot AFB, Offutt AFB, and Scott AFB. These six were the western most locations in the research. Also for Minot AFB, no decision ruled allowed for the forecast of CG lightning in the months of March and April. After reviewing the data it was determined this was not an error. In 11 years there were zero occurrences of CG lighting within 25nm for the month of March. There were only eight total Cg lightning occurrences within 25 nm in April. The use of month as a predictor proved that the time of year is critical for this region. Also, the Minot AFB decision tool cannot be used to forecast CG lightning in March or April. Further investigations could be done on these four locations to see if a month dependent list of stability indices could be created, however, that was beyond the scope of this research. The forecast rules were extracted from the final 52 decision trees and placed into tables (see Appendix B).

The results achieved were higher than anticipated in most aspects. The overall KSS score was from 0.31 to 0.70 and this showed the decision trees had some improvement over random forecasts. Also, the POD was over 70% in 48 of 52 decision trees. However, as positive as those results were, the FAR was extremely high ranging from 0.43 to as high 0.89. Also, all bias scores were above one for every decision tree which showed the decision trees had a tendency to over forecasted CG lightning. The entire results of the statistical verification test can be found in Appendix A.

b. Recommendations

The decision trees developed by CART provided valuable insight in the forecasting of weather parameters and could provide a possible avenue for future research. Any weather phenomena can theoretically be predicted if the right predictors are available. As CART becomes a more popular and proven tool, more research should be done to test further CART applications into weather forecasting. Future CART decision trees should use the guidance and shortcomings provided by this and other previous research to further enhance the predictive outcome of weather decision models.

Although the results in this research provided promising values for a CG lightning decision tool, it is not recommended to use these decision tools as a stand alone CG lightning forecast tool. It is also not recommended to apply these rules without using the specific sites paired with the correct upper air site. A problem with this tool was it predicts the probability that CG lightning will occur in the next 12 hours and not a specific time period within the 12 hours. The tool does not account for any parameters except the ones used in the decision tree, so this tool should never be used alone in the decision process of forecasting CG lightning. The forecast rules should be used in conjunction with forecaster knowledge, numerical weather prediction models and other tools to ensure an accurate forecast.

A major shortfall of the method used in this research was the 12 hour time blocks of data. The prediction rates are based on the ability of CART to correctly classify the occurrence of CG lightning anytime during the entire 12 hour period. One recommendation for further research is to use hourly surface data and CG lightning data to develop forecast aides that have a stronger impact on weather forecasts.

Another recommendation would be to use model data combined with CART to predict CG lightning. The same procedures could be followed except replacing upper air stability indices and other predictors with predictors from the model. For example, use the current surface temperature, 3-hour forecast temperature and 3-hour dew point as predictors to develop a better decision tool. This could allow for three, six, nine, 12-

hour, or even longer CG lightning forecast. This would help eliminate the uncertainty in occurrence associated with the method in this research.

Appendix A: Two by Two Matrices with Results

This appendix presents the complete list of two by two matrices and the calculated verification test results. The CART analysis program was used to data mine for the 13 locations in the 15 AOR. The locations were combined with a specific upper air station that was considered representative of the location (see table 12 for the correct station). The data used in the building of the model and the resulting decision trees were from 1993 to 2002. Verification of the model and resulting decision trees was accomplished with independent data from 2003. This method was chosen to test the benefits as a stand along forecast tool. Once in place in the field, the decision aide will not be update annually so it was important to verify it as forecast tool and not just a research tool.

The contingency tables were created to display the results from the CART data analysis. Verification test were calculated on the contingency tables and all were displayed as a single results table. Each table contains an observed versus forecast two by two matrix, the hit rate (HR), a threat score (TS), the probability of detection (POD), the false-alarm rate (FAR), the bias, and the Kuipers skill score (KSS) (Wilks 1995). For further information and the equations used for the calculations of the tests, see chapter 3 section c.

Eight result tables are provided for each location: 1) model results for a 10 nm radius with a one to one cost, 2) verification results for a 10 nm radius with a one to one cost, 3) model results for a 10 nm radius with a two to one cost, 4) verification results for a 10 nm radius with a two to one cost 5) model results for a 25 nm radius with a one to one cost, 6) verification results for a 25 nm radius with a one to one cost, 7) model results for a 25 nm radius with a two to one cost, and 8) verification results for a 25 nm radius with a two to one cost.

A1. Andrews AFB

Table A1. Model results for Andrews AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	371	1343	66%	0.20	0.78	0.78	3.6	0.42
NO	106	2460						

Table A2. Verification results for Andrews AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	63	124	69%	0.32	0.90	0.66	2.7	0.55
NO	7	234						

Table A3. Model results for Andrews AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	412	1785	57%	0.18	0.96	0.81	4.6	0.39
NO	65	2018						

Table A4. Verification results for Andrews AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	68	163	61%	0.29	0.97	0.71	3.3	0.52
NO	2	195						

Table A5. Model results for Andrews AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	652	1000	73%	0.36	0.79	0.61	2.0	0.50
NO	170	2458						

Table A6. Verification results for Andrews AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	100	74	80%	0.54	0.90	0.43	1.6	0.67
NO	11	243						

Table A7. Model results for Andrews AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	736	1598	61%	0.30	0.89	0.68	2.8	0.43
NO	86	1860						

Table A8. Verification results for Andrews AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	109	142	66%	0.43	0.98	0.57	2.3	0.53
NO	2	175						

A2. Dover AFB

Table A9. Model results for Dover AFB 10 nm (cost one to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	350	1292	68%	0.20	0.81	0.79	3.8	0.47
NO	82	2556						

Table A10. Verification results for Dover AFB 10 nm (cost one to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	59	111	72%	0.33	0.89	0.65	2.6	0.59
NO	7	251						

Table A11. Model results for Dover AFB 10 nm (cost two to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	394	1780	58%	0.18	0.91	0.82	5.0	0.45
NO	38	2068						

Table A12. Verification results for Dover AFB 10 nm (cost two to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	64	168	60%	0.27	0.97	0.72	3.5	0.51
NO	2	194						

Table A13. Model results for Dover AFB 25 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	632	1171	70%	0.33	0.83	0.65	2.4	0.50
NO	129	2348						

Table A14. Verification results for Dover AFB 25 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	93	99	75%	0.46	0.90	0.52	1.9	0.60
NO	10	226						

Table A15. Model results for Dover AFB 25 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	670	1501	63%	0.30	0.88	0.69	2.9	0.45
NO	91	2018						

Table A16. Verification results for Dover AFB 25 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	96	136	67%	0.40	0.93	0.59	2.3	0.51
NO	7	189						

A3. Ellsworth AFB

Table A17. Model results for Ellsworth AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	463	1120	71%	0.27	0.78	0.71	2.7	0.48
NO	131	2566						

Table A18. Verification results for Ellsworth AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	40	116	68%	0.22	0.63	0.74	2.5	0.32
NO	23	249						

Table A19. Model results for Ellsworth AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	564	1558	63%	0.26	0.95	0.55	3.6	0.53
NO	30	2128						

Table A20. Verification results for Ellsworth AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	50	178	55%	0.21	0.79	0.78	3.6	0.31
NO	13	187						

Table A21. Model results for Ellsworth AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	882	1053	71%	0.41	0.81	0.54	1.8	0.48
NO	205	2140						

Table A22. Verification results for Ellsworth AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	88	102	69%	0.40	0.75	0.54	1.6	0.42
NO	29	209						

Table A23. Model results for Ellsworth AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	1003	1300	68%	0.42	0.92	0.56	2.1	0.52
NO	84	1893						

Table A24. Verification results for Ellsworth AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	103	130	66%	0.42	0.88	0.56	2.0	0.46
NO	14	181						

A4. Ft Belvoir

Table A25. Model results for Ft Belvoir 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	413	1358	66%	0.22	0.82	0.77	3.5	0.46
NO	90	2419						

Table A26. Verification results for Ft Belvoir 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	66	132	68%	0.33	0.94	0.67	2.8	0.57
NO	4	226						

Table A27. Model results for Ft Belvoir 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	427	1720	58%	0.19	0.85	0.80	4.3	0.39
NO	76	2057						

Table A28. Verification results for Ft Belvoir 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	69	159	63%	0.30	0.99	0.70	3.3	0.54
NO	1	199						

Table A29. Model results for Ft Belvoir 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	670	1056	71%	0.35	0.79	0.61	2.0	0.48
NO	181	2373						

Table A30. Verification results for Ft Belvoir 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	101	99	74%	0.47	0.88	0.50	1.7	0.56
NO	14	214						

Table A31. Model results for Ft Belvoir 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	729	1445	63%	0.32	0.86	0.66	2.6	0.44
NO	122	1984						

Table A32. Verification results for Ft Belvoir 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	107	125	69%	0.45	0.93	0.54	2.0	0.53
NO	8	188						

A5. Ft Drum

Table A33. Model results for Ft Drum 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	300	1212	71%	0.19	0.87	0.80	4.4	0.56
NO	45	2723						

Table A34. Verification results for Ft Drum 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	24	110	73%	0.17	0.86	0.89	4.8	0.58
NO	4	290						

Table A35. Model results for Ft Drum 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	327	1383	67%	0.19	0.95	0.81	5.0	0.60
NO	18	2552						

Table A36. Verification results for Ft Drum 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	25	148	65%	0.14	0.89	0.86	6.2	0.52
NO	3	252						

Table A37. Model results for Ft Drum 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	584	892	78%	0.38	0.93	0.60	2.3	0.68
NO	46	2758						

Table A38. Verification results for Ft Drum 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	42	94	75%	0.28	0.76	0.69	2.5	0.51
NO	13	279						

Table A39. Model results for Ft Drum 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	578	1500	64%	0.27	0.92	0.72	3.3	0.51
NO	52	2150						

Table A40. Verification results for Ft Drum 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	50	154	63%	0.24	0.91	0.75	3.7	0.50
NO	5	219						

A6. Grand Forks

Table A41. Model results for Grand Forks AFB 10 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	401	1271	70%	0.24	0.93	0.76	3.9	0.60
NO	30	2578						

Table A42. Verification results for Grand Forks AFB 10 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	31	117	71%	0.20	0.78	0.79	3.7	0.47
NO	9	271						

Table A43. Model results for Grand Forks AFB 10 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	397	1372	67%	0.22	0.92	0.78	4.1	0.56
NO	34	2477						

Table A44. Verification results for Grand Forks AFB 10 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	32	133	67%	0.18	0.80	0.81	4.1	0.46
NO	8	255						

Table A45. Model results for Grand Forks AFB 25 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	608	1028	73%	0.35	0.84	0.63	2.3	0.55
NO	118	2526						

Table A46. Verification results for Grand Forks AFB 25 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	46	100	66%	0.28	0.71	0.68	2.2	0.43
NO	19	263						

Table A47. Model results for Grand Forks AFB 25 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	706	1842	56%	0.27	0.97	0.72	3.5	0.45
NO	20	1712						

Table A48. Verification results for Grand Forks AFB 25 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	64	188	56%	0.25	0.98	0.75	3.9	0.47
NO	1	175						

A7. Grissom AFB

Table A49. Model results for Grissom AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	506	1469	63%	0.24	0.81	0.74	3.2	0.41
NO	120	2185						

Table A50. Verification results for Grissom AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	42	102	70%	0.24	0.60	0.71	2.1	0.32
NO	28	256						

Table A51. Model results for Grissom AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	599	1898	55%	0.24	0.96	0.76	4.0	0.44
NO	27	1756						

Table A52. Verification results for Grissom AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	58	152	62%	0.26	0.83	0.72	3.0	0.40
NO	12	206						

Table A53. Model results for Grissom AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	807	1488	61%	0.33	0.83	0.65	2.4	0.38
NO	163	1822						

Table A54. Verification results for Grissom AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	72	93	70%	0.36	0.67	0.56	1.5	0.38
NO	35	228						

Table A55. Model results for Grissom AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	910	1931	53%	0.31	0.94	0.68	2.9	0.35
NO	60	1379						

Table A56. Verification results for Grissom AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	92	151	61%	0.36	0.86	0.62	2.3	0.39
NO	15	170						

A8. McGuire AFB

Table A57. Model results for McGuire AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	323	1421	64%	0.17	0.76	0.81	4.1	0.39
NO	102	2434						

Table A58. Verification results for McGuire AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	38	145	64%	0.20	0.81	0.79	3.9	0.43
NO	9	236						

Table A59. Model results for McGuire AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	400	1965	54%	0.17	0.94	0.83	5.6	0.43
NO	25	1890						

Table A60. Verification results for McGuire AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	46	216	49%	0.17	0.98	0.82	5.6	0.41
NO	1	165						

Table A61. Model results for McGuire AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	606	1318	66%	0.29	0.82	0.69	2.6	0.44
NO	136	2220						

Table A62. Verification results for McGuire AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	70	129	66%	0.33	0.81	0.65	2.3	0.44
NO	16	213						

Table A63. Model results for McGuire AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	695	1769	58%	0.28	0.94	0.72	3.3	0.44
NO	47	1769						

Table A64. Verification results for McGuire AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	84	181	57%	0.31	0.98	0.68	3.1	0.45
NO	2	161						

A9. Minot AFB

Table A65. Model results for Minot AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	410	911	78%	0.30	0.94	0.69	3.0	0.70
NO	26	2933						

Table A66. Verification results for Minot AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	20	91	76%	0.16	0.63	0.82	3.5	0.40
NO	12	305						

Table A67. Model results for Minot AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	418	1654	61%	0.20	0.96	0.80	4.8	0.53
NO	18	2190						

Table A68. Verification results for Minot AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	27	161	61%	0.14	0.84	0.86	5.9	0.44
NO	5	235						

Table A69. Model results for Minot AFB 25 nm (cost one to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	612	1096	72%	0.34	0.84	0.64	2.4	0.53
NO	114	2458						

Table A70. Verification results for Minot AFB 25 nm (cost one to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	45	118	69%	0.25	0.76	0.72	2.8	0.44
NO	14	251						

Table A71. Model results for Minot AFB 25 nm (cost two to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	690	1394	67%	0.33	0.95	0.67	2.9	0.56
NO	36	2160						

Table A72. Verification results for Minot AFB 25 nm (cost two to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	46	147	63%	0.22	0.78	0.76	3.3	0.38
NO	13	222						

A10. Offutt AFB

Table A73. Model results for Offutt AFB 10 nm (cost one to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	499	1524	62%	0.23	0.81	0.75	3.3	0.40
NO	115	2142						

Table A74. Verification results for Offutt AFB 10 nm (cost one to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	52	164	60%	0.23	0.87	0.76	3.6	0.42
NO	8	204						

Table A75. Model results for Offutt AFB 10 nm (cost two to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	579	1807	57%	0.24	0.94	0.76	3.9	0.45
NO	35	1859						

Table A76. Verification results for Offutt AFB 10 nm (cost two to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	55	191	54%	0.22	0.92	0.78	4.1	0.40
NO	5	177						

Table A77. Model results for Offutt AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	884	1034	73%	0.43	0.86	0.54	1.9	0.55
NO	140	2222						

Table A78. Verification results for Offutt AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	77	107	69%	0.37	0.76	0.58	1.8	0.44
NO	24	220						

Table A79. Model results for Offutt AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	982	1380	67%	0.41	0.96	0.58	2.3	0.54
NO	42	1876						

Table A80. Verification results for Offutt AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	87	150	62%	0.35	0.86	0.63	2.3	0.40
NO	14	177						

A11. Scott AFB

Table A81. Model results for Scott AFB 10 nm (cost one to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	527	1192	68%	0.28	0.75	0.69	2.4	0.42
NO	176	2385						

Table A82. Verification results for Scott AFB 10 nm (cost one to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	51	132	66%	0.26	0.78	0.72	2.8	0.42
NO	14	231						

Table A83. Model results for Scott AFB 10 nm (cost two to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	677	1757	58%	0.28	0.96	0.72	3.5	0.47
NO	26	1820						

Table A84. Verification results for Scott AFB 10 nm (cost two to one)

Forecasted	**Observed**		**HR**	**TS**	**POD**	**F A R**	**Bias**	**KSS**
	YES	**NO**						
YES	57	184	55%	0.23	0.88	0.76	3.7	0.37
NO	8	179						

Table A85. Model results for Scott AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	954	1072	72%	0.44	0.88	0.53	1.9	0.54
NO	136	2118						

Table A86. Verification results for Scott AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	73	117	66%	0.34	0.73	0.62	1.9	0.37
NO	27	211						

Table A87. Model results for Scott AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	1061	1314	69%	0.44	0.97	0.55	2.2	0.56
NO	29	1876						

Table A88. Verification results for Scott AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	84	150	61%	0.34	0.84	0.64	2.3	0.38
NO	16	178						

A12. Westover AFB

Table A89. Model results for Westover AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	302	1275	69%	0.19	0.89	0.81	4.7	0.57
NO	37	2666						

Table A90. Verification results for Westover AFB 10 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	23	143	66%	0.14	0.88	0.86	6.4	0.53
NO	3	259						

Table A91. Model results for Westover AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	334	1284	70%	0.21	0.99	0.79	4.8	0.66
NO	5	2657						

Table A92. Verification results for Westover AFB 10 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	22	149	64%	0.13	0.85	0.87	6.6	0.48
NO	4	253						

Table A93. Model results for Westover AFB 25 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	525	1052	74%	0.32	0.87	0.67	2.6	0.59
NO	77	2626						

Table A94. Verification results for Westover AFB 25 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	47	125	70%	0.27	0.92	0.73	3.4	0.59
NO	4	252						

Table A95. Model results for Westover AFB 25 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	567	1248	70%	0.31	0.94	0.69	3.0	0.60
NO	35	2430						

Table A96. Verification results for Westover AFB 25 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	47	148	64%	0.24	0.92	0.76	3.8	0.53
NO	4	229						

A13. Wright Patterson AFB

Table A97. Model results for Wright Patterson AFB 10 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	589	1211	69%	0.31	0.85	0.67	2.6	0.51
NO	107	2373						

Table A98. Verification results for Wright Patterson AFB 10 nm (cost one to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	75	113	70%	0.37	0.84	0.60	2.1	0.51
NO	14	226						

Table A99. Model results for Wright Patterson AFB 10 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	657	1754	58%	0.27	0.94	0.73	3.5	0.45
NO	39	1830						

Table A100. Verification results for Wright Patterson AFB 10 nm (cost two to one)

Forecasted	Observed YES	Observed NO	HR	TS	POD	F A R	Bias	KSS
YES	85	164	61%	0.34	0.96	0.66	2.8	0.47
NO	4	175						

Table A101. Model results for Wright Patterson AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	844	873	74%	0.43	0.79	0.51	1.6	0.52
NO	224	2339						

Table A102. Verification results for Wright Patterson AFB 25 nm (cost one to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	107	107	77%	0.46	0.84	0.50	1.7	0.51
NO	21	224						

Table A103. Model results for Wright Patterson AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	1009	1558	62%	0.38	0.94	0.61	2.4	0.46
NO	59	1654						

Table A104. Verification results for Wright Patterson AFB 25 nm (cost two to one)

Forecasted	Observed		HR	TS	POD	F A R	Bias	KSS
	YES	NO						
YES	124	143	66%	0.46	0.97	0.54	2.1	0.49
NO	4	157						

Appendix B: Forecast Rules

This appendix presents the complete list of forecast rules for the 13 locations in the 15^{th} OWS AOR. The tables were composed from the positive CG lightning results found using the CART analysis. The positive terminal nodes for the learning tree were required to have at least a 1% total probability to be included in the set of rules. All terminals nodes not meeting this requirement were eliminated. These forecast rules were created using data from a specific upper air station (see table 12 for the correct station). The learning data is from 1993 to 2002 and was used to build the models. Independent test data from 2003 was used to validate the models.

The cost matrix used to create each decision tree was based on the importance of correctly forecasting CG lightning. A one to one cost weighed the importance of correctly forecasting CG lightning and false alarm rates equally. A tree built using a two to one cost matrix expressed the importance of correctly forecasting CG lightning. These trees simply implied that it was twice as important to not miss a CG lightning occurrence as it was to forecast for CG lightning and it not occur.

Four tables are provided for each location: 1) a 10 nm radius with a one to one cost, 2) a 10 nm radius with a two to one cost, 3) a 25 nm radius with a one to one cost, and 4) a 25 nm radius with a two to one cost. Each table contains forecast rules, and the probabilities of forecasting CG lightning correctly. This probability is calculated using Bayes Rules (see chapter 4).

Tables B-1 through B-4 are the forecast rules and probabilities for Andrews Air Force Base, Maryland. Washington/ Dules, VA (724030) was the upper air site matched with Andrews. The following combination of CG lightning predictors were determined to produce the best success rate for Andrews: 1) Showalters Stability Index (SSI), 2) K Index (KI), 3) Severe Weather Threat Index (SWEAT), 4) Lifted Index (LI), and 5) Surface pressure at Andrews (STAPRESS).

Table B-1. Rules to predict CG lightning for Andrews AFB within 10 nm (cost one to one)

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 0.04	0.79
IF SSI > 0.04 AND SSI <= 3.66 AND KI <= 28.95 AND SWEAT > 206.5	0.62
IF SSI > 0.04 AND SSI <= 3.66 AND KI > 28.95	0.61

Table B-2. Rules to predict CG lightning for Andrews AFB within 10 nm (cost two to one)

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.66	0.65

Table B-3. Rules to predict CG lightning for Andrews AFB within 25 nm (cost one to one)

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.76 AND KI <= 28.85 AND STAPRESS <= 1010.2 AND LI <= -0.85	0.69
IF SSI <= 3.76 AND KI <= 28.85 AND LI > -0.85 AND STAPRESS <= 998.0	0.76
IF SSI <= 3.76 AND STAPRESS > 998.0 AND STAPRESS <= 1010.2 AND KI > 23.45 AND KI <= 28.85 AND LI > -0.85 AND LI <= 3.95	0.62
IF SSI <= 3.76 AND KI > 28.85 AND LI <= 0.25	0.76
IF SSI <= 3.76 AND KI > 28.85 AND LI > 0.25 AND STAPRESS <= 1008.4	0.71

Table B-4. Rules to predict CG lightning for Andrews AFB within 25 nm (cost two to one)

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.76	0.66
IF SSI > 3.76 AND STAPRESS <= 1007.72 AND KI > 12.85 AND LI <= 4.35	0.54

Tables B-5 through B-8 are the forecast rules and probabilities for Dover Air Force Base, Delaware. Washington/ Dules, VA (724030) was the upper air site matched with Dover. The following combination of CG lightning predictors were determined to produce the best success rate for Dover: 1) Showalters Stability Index (SSI), 2) Convective Available Potential Energy (CAPE), and 3) Total Totals Index (TTI).

Table B-5. Rules to predict CG lightning for Dover AFB within 10 nm (cost one to one).

Rules to Forecast CG lightning	**Probability of correctly predicting CG lightning**
IF SSI <= 3.74 AND CAPE > 177.1	0.71

Table B-6. Rules to predict CG lightning for Dover AFB within 10 nm (cost two to one).

Rules to Forecast CG lightning	**Probability of correctly predicting CG lightning**
IF SSI <= 3.74	0.66

Table B-7. Rules to predict CG lightning for Dover AFB within 25 nm (cost one to one).

Rules to Forecast CG lightning	**Probability of correctly predicting CG lightning**
IF SSI <= 1.46	0.72
IF SSI > 1.46 AND SSI <= 3.52 AND CAPE <= 303.3 AND TTI > 47.05	0.62
IF SSI > 1.46 AND SSI <= 3.52 AND TTI > 40.55 AND CAPE > 303.3	0.68

Table B-8. Rules to predict CG lightning for Dover AFB within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.74	0.67

Tables B-9 through B-12 are the forecast rules and probabilities for Ellsworth Air Force Base, South Dakota. Rapid City Regional Airport (726620) was the upper air site matched with Ellsworth AFB. The following combination of CG lightning predictors were determined to produce the best success rate for Ellsworth: 1) Lifted Index (LI), 2) Showalters Stability Index (SSI), 3) Total Totals Index (TTI) 4) K Index (KI), 5) Month, 6) KO Index (KO), 7) Severe Weather Threat Index (SWEAT), 8) Surface pressure at Ellsworth (STAPRESS), 9) Surface pressure at Rapid City Regional Airport (PRESS_UA), and 10) Three hour surface pressure change at Ellsworth matching the sounding time (PRESC3).

Table B-9. Rules to predict CG lightning for Ellsworth AFB within 10 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF LI <= 2.65 AND SSI <= -1.19	0.85
IF LI <= 2.65 AND SSI > -1.19 AND TTI <= 53.5 AND KI > 25.05	0.62

Table B-10. Rules to predict CG lightning for Ellsworth AFB within 10 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF LI <= -0.45	0.75
IF MONTH = 3 OR 4 AND LI > -0.45 AND LI <= 3.65 AND KO > 1.3 AND SWEAT <= 113.5	0.58
IF MONTH = 5, 6, 7, 8, OR 9 AND LI > -0.45 AND LI <= 2.65 AND STAPRESS > 889.7 AND STAPRESS <= 909.7 AND KO > -7.95 AND KO <= 0.55	0.58
IF MONTH = 5, 6, 7, 8, OR 9 AND LI > -0.45 AND LI <= 3.65 AND KO > 0.55 AND PRESS_UA <= 907.5	0.71
IF MONTH = 3, 4, 5, OR 9 AND LI > 3.65 AND SWEAT > 142 AND KI > 16.2	0.55
IF MONTH = 6, 7, or 8 AND LI > 3.65 AND SSI <= 6.38 AND KO <= 6.65 AND TTI <= 41.5 AND PRESS_UA <= 903.5 AND PRESC3 > -1.1	0.54

Table B-11. Rules to predict CG lightning for Ellsworth AFB within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= -1.17	0.83
IF MONTH = 5, 6, 7, 8, OR 9 AND SSI > -1.17 AND SSI <= 2.90	0.64

Table B-12. Rules to predict CG lightning for Ellsworth AFB within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF MONTH = 5, 6, 7, 8, OR 9 AND SSI <= 3.00 AND LI <= 0.25	0.76
IF MONTH = 5, 6, 7, 8, OR 9 AND SSI <= 3.00 AND LI > 0.25 AND KO > -6.35	0.64
IF MONTH = 8 AND SSI > 3.00 AND SSI <= 6.93 AND KI <= 23.95 AND KO <= 4.25	0.44
IF SSI > 3.00 AND SSI <= 6.93 AND KI <= 23.95 AND KO > 4.25 AND CAPE > 7.4	0.57
IF MONTH = 6 or 7 AND SSI > 3.00 AND SSI <= 6.93 AND KI > 23.95 AND KO <= 6.65	0.54
IF SSI > 3.00 AND SSI <= 6.93 AND KI > 23.95 AND KO > 6.65	0.72

Tables B-13 through B-16 are the forecast rules and probabilities for Fort Belvoir, Virginia. Washington/Dulles (724030) was the upper air site matched with Ft Belvoir. The following CG lightning predictors were determined to produce the best success rate for Ft Belvoir: 1) Lifted Index (LI), 2) Showalters Stability Index (SSI), 3) K Index (KI), 5) Severe Weather Threat Index (SWEAT), and 8) Convective Available Potential Energy (CAPE).

Table B-13. Rules to predict CG lightning for Ft Belvoir within 10 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF KI <= 26.65 AND SSI <= 1.03 AND LI <= -0.95	0.65
IF KI <= 26.65 AND SSI > 1.03 AND SSI <= 1.28	0.78
SSI <= 1.28 AND KI > 26.65	0.72
IF SSI > 1.28 AND SSI <= 3.61 AND SWEAT > 132.5 AND CAPE > 132.9 AND KI <= 16.6	0.75
IF SSI > 1.28 AND SSI <= 3.61 AND SWEAT > 132.5 AND KI > 24.25	0.57

Table B-14. Rules to predict CG lightning for Ft Belvoir within 10 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.61	0.65

Table B-15. Rules to predict CG lightning for Ft Belvoir within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.74 AND KI <= 28.85 AND CAPE > 82.8 AND SWEAT <= 207.5 AND LI <= -1.25	0.63
IF SSI <= 3.74 AND KI <= 28.85 AND SWEAT <= 207.5 AND LI > -1.25 AND CAPE > 82.8 AND CAPE <= 209.5	0.64
IF SSI <= 3.74 AND KI <= 28.85 AND CAPE > 82.8 AND SWEAT > 207.5	0.68
IF SSI <= 3.74 AND KI > 28.85	0.73

Table B-16. Rules to predict CG lightning for Ft Belvoir within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.74	0.67

Tables B-17 through B-20 are the forecast rules and probabilities for Fort Drum, New York. Buffalo International Airport (725280) was the upper air site matched with Ft Drum. The following CG lightning predictors were determined to produce the best success rate for Ft Drum: 1) Lifted Index (LI), 2) Showalters Stability Index (SSI), 3) Total Totals Index (TTI) 4) K Index (KI), 5) Convective Available Potential Energy (CAPE), 6) KO Index (KO), 7) Severe Weather Threat Index (SWEAT), and 8) Temperature in Celsius for Ft Drum at sounding time (TEMPC).

Table B-17. Rules to predict CG lightning for Ft Drum within 10 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF CAPE <= 160.6 AND KI > 24.15 AND SWEAT > 192.5 AND TEMPC <= 16.6	0.65
IF CAPE > 160.6	0.78

Table B-18. Rules to predict CG lightning for Ft Drum within 10 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF CAPE <= 160.6 AND KI > 24.45 AND SWEAT <= 192.5 AND SSI > 1.74 AND SSI <= 4.77 AND TTI > 43.1	0.57
IF SSI <= 4.77 AND CAPE <= 160.6 AND KI > 24.45 AND SWEAT > 192.5	0.68
IF SSI <= 4.77 AND CAPE > 160.6	0.75
IF SWEAT > 66.5 AND SSI > 8.60 AND LI <= 6.95	0.73
IF SSI > 8.60 AND LI > 6.95 AND LI <= 10.45 AND SWEAT > 105.5	0.45

Table B-19. Rules to predict CG lightning for Ft Drum within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF LI <= 2.35 AND CAPE > 183.6 AND KI > 12.4	0.83
IF LI > 2.35 AND LI <= 6.95 AND KI > 21.05 AND SWEAT > 185.5 AND SWEAT <= 214.5	0.75
IF LI > 2.35 AND LI <= 6.95 AND SWEAT > 214.5 AND SSI > 2.31 AND KI > 21.05 AND KI <= 30.85 AND TTI > 38.7	0.69
IF LI > 2.35 AND LI > 6.95 AND SWEAT <= 123.5 AND MONTH = 5, 6, 7, 8, OR 9 AND SWEAT > 63.5 AND SSI <= 10.40 AND TTI <= 41.15 AND SWEAT > 72.5 AND TTI <= 36.95 AND SWEAT > 98	0.64

Table B-20. Rules to predict CG lightning for Ft Drum within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF LI <= 6.05 AND CAPE <= 138.1 AND KI > -3.6 AND KI <= 27.05 AND KO > -5.75 AND KO <= 3.65	0.39
IF LI <= 6.05 AND CAPE <= 138.1 AND KO > -5.75 AND KI > 27.05	0.59
IF LI <= 6.05 AND CAPE > 138.1	0.75

Tables B-21 through B-24 are the forecast rules and probabilities for Grand Forks Air Force Base, North Dakota. Bismarck Municipal Airport (7276840) was the upper air site matched with Grand Forks. The following CG lightning predictors were determined to produce the best success rate for grand Forks: 1) Lifted Index (LI), 2) Showalters Stability Index (SSI), 3) Total Totals Index (TTI), 4) K Index (KI), 5) Month, 6) Convective Available Potential Energy (CAPE), and 7) Severe Weather Threat Index (SWEAT).

Table B-21. Rules to predict CG lightning for Grand Forks within 10 nm (cost one to one).

Rules to Forecast CG lightning	**Probability of correctly predicting CG lightning**
IF MONTH = 5, 6, 7, 8, OR 9 AND SSI <= 2.66	0.75
IF MONTH = 4, 6, 7, OR 8 AND SSI > 2.66 AND SSI <= 5.90 AND KI > 17	0.60
IF SSI > 2.66 AND SSI <= 5.90 AND CAPE > 728.9	0.84

Table B-22. Rules to predict CG lightning for Grand Forks within 10 nm (cost two to one).

Rules to Forecast CG lightning	**Probability of correctly predicting CG lightning**
IF SSI <= 2.27	0.73
IF SSI > 2.27 AND SSI <= 4.58 AND CAPE <= 10.5 AND LI <= 5.45	0.53
IF CAPE > 10.5 AND CAPE <= 715.7 AND TTI <= 45.7 AND SSI > 2.27 AND SSI <= 2.89	0.60
IF SSI > 2.27 AND SSI <= 4.58 AND CAPE > 715.7	0.83

Table B-23. Rules to predict CG lightning for Grand Forks within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF LI <= 4.75 AND SSI <= 1.32	0.77
IF SSI > 1.32 AND KI <= 25.25 AND CAPE <= 65.9 AND LI <= 3.05	0.63
IF LI <= 4.75 AND SSI > 1.32 AND KI > 25.25	0.64

Table B-24. Rules to predict CG lightning for Grand Forks within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF LI <= 4.75	0.69
IF LI > 4.75 AND KI > 14.65 AND SWEAT > 80.5	0.36

Tables B-25 through B-28 are the forecast rules and probabilities for Grissom Air Force Base, Indiana. Davenport Upper Air (744550) was the upper air site matched with Grissom. The following CG lightning predictors were determined to produce the best success rate for Grissom: 1) Showalters Stability Index (SSI), 3) Total Totals Index (TTI) 4) K Index (KI), 5) Month, 7) Severe Weather Threat Index (SWEAT), 8) Surface pressure at Davenport (PRESS_UA), and 10) Temperature in Celsius for Grissom at sounding time (TEMPC).

Table B-25. Rules to predict CG lightning for Grissom within 10 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF MONTH = 4, 5, 6, 7, 8, OR 9 AND SSI <= 3.93 AND TEMPC > 1.4 AND KI > 19.85 & KI <= 30.15	0.54
IF SSI <= 3.93125 AND KI > 30.15	0.75

Table B-26. Rules to predict CG lightning for Grissom within 10 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF MONTH = 4, 5, 7, 8, OR 9 AND SSI <= 3.93 AND TEMPC > 0.3 AND KI <= 19.85 AND TTI > 43.0	0.54
IF MONTH = 4, 5, 6, 7, 8, OR 9 AND SSI <= 3.93 AND TEMPC > 0.3 AND KI > 19.85 AND KI <= 30.15	0.53
IF SSI <= 3.93 AND KI > 30.15	0.75
IF MONTH = 3, 7, 8, OR 9 AND SSI > 3.93125 AND KI > 0 AND TEMPC > 25.3	0.48
IF MONTH = 4, 5, OR 6 AND KI > 0 AND PRESS_UA <= 987.1 AND SSI > 3.93 AND SSI <= 6.75	0.62
IF MONTH = 4, 5, OR 6 AND PRESS_UA <= 987.1 AND SSI > 6.75 AND KI > 11.8	0.50
IF MONTH = 4, 5, OR 6 AND SSI > 3.93 AND KI > 0 AND PRESS_UA > 987.1 AND TEMPC <= 19.5 AND SWEAT > 124.5	0.55

Table B-27. Rules to predict CG lightning for Grissom within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF KI <= 23.95 AND SSI <= 0.33	0.70
IF SSI <= 3.98 AND KI > 23.95	0.65
IF SSI > 3.98 AND KI > 16.65 AND SWEAT > 145.5	0.65

Table B-28. Rules to predict CG lightning for Grissom within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 4.65	0.62
IF SWEAT > 58.5 AND SWEAT <= 117.5 AND KI > 21.15 AND SSI > 5.55	0.63
IF SSI > 4.65 AND KI > 1.85 AND SWEAT > 145.5	0.49

Tables B-29 through B-32 are the forecast rules and probabilities for McGuire Air Force Base, New Jersey. Pittsburgh (725200) was the upper air site matched with McGuire. The following CG lightning predictors were determined to produce the best success rate for McGuire: 1) Lifted Index (LI), 2) Showalters Stability Index (SSI), 3) KO Index (KO), 4) K Index (KI), 6) Convective Available Potential Energy (CAPE), and 7) Severe Weather Threat Index (SWEAT).

Table B-29. Rules to predict CG lightning for McGuire within 10 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF CAPE > 94.0	0.67

Table B-30. Rules to predict CG lightning for McGuire within 10 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF CAPE <= 33.4 AND KI > 23.45 AND SSI > 2.03	0.53
IF CAPE > 33.4 AND KO <= 4.75	0.65

Table B-31. Rules to predict CG lightning for McGuire within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF CAPE <= 46.5 AND SSI > 1.95 AND KI > 21.55 AND KI <= 26.75 AND KO > 10.1	0.64
IF CAPE <= 46.5 AND KO > -8.45 AND SSI > 1.95 AND KI > 26.75	0.61
IF KI <= 21.75 AND KO > -13.85 AND KO <= -2.65 AND CAPE > 183.0	0.61
IF CAPE > 46.5 AND KI > 21.75 AND KO <= 4.75	0.70

Table B-32. Rules to predict CG lightning for McGuire within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 5.21 AND CAPE <= 38.8 AND SWEAT <= 152.5 AND KI > 19.4	0.43
IF SSI <= 5.21 AND CAPE <= 38.8 AND SWEAT > 152.5 AND SWEAT <= 288.5	0.55
IF SSI <= 5.21 AND CAPE > 38.8 AND SWEAT <= 150.5 AND KI > 22.45	0.60
IF SSI <= 5.21 AND CAPE > 38.8 AND SWEAT > 150.5	0.69
IF SWEAT > 80.5 AND CAPE <= 147 AND KI <= 21.55 AND SSI > 5.21 AND SSI <= 17.23 AND LI > 6.75 AND LI <= 9.15	0.41
IF SWEAT > 80.5 AND CAPE <= 147 AND KI > 21.55 AND SSI > 6.25	0.63
IF SSI > 5.21 AND SWEAT > 80.5 AND CAPE > 147	0.70

Tables B-33 through B-36 are the forecast rules and probabilities for Minot Air Force Base, North Dakota. Mount Glasgow International Airport (727680) was the upper air site matched with Minot. The following CG lightning predictors were determined to produce the best success rate for Minot: 1) Lifted Index (LI), 2) Showalters Stability Index (SSI), 3) Total Totals Index (TTI) 4) K Index (KI), 5) Month, 6) Convective Available Potential Energy (CAPE), 7) Severe Weather Threat Index (SWEAT), 8) Surface pressure at Mt Glasgow Intl Airport (PRESS_UA), and 9) 3 hour surface pressure change at Minot matching the sounding time (PRESC3).

Table B-33. Rules to predict CG lightning for Minot within 10 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF MONTH = 6, 7, OR 8 AND PRESS_UA <= 934.5 AND KI <= 16.35 AND CAPE > 0.5 AND CAPE <= 1049.1	0.69
IF MONTH = 5, 6, 7, 8, OR 9 AND KI > 21.85 AND LI <= -0.55 AND TTI > 47.7	0.82
IF MONTH = 5, 6, 7, 8, OR 9 AND LI > -0.55 AND LI <= 2.75 AND KI > 21.85 AND KI <= 25.95 AND CAPE <= 154.9 AND PRESS_UA <= 936.5 AND TTI <= 51.4 AND SWEAT <= 74	0.89
IF MONTH = 5, 6, 7, 8, OR 9 AND LI > -0.55 AND LI <= 2.75 AND KI > 21.85 AND KI <= 25.95 AND CAPE <= 154.9 AND PRESS_UA <= 936.5 AND TTI <= 51.4 AND SWEAT > 97.5	0.73
IF MONTH = 5, 6, 7, 8, OR 9 AND LI > -0.55 AND LI <= 2.75 AND KI > 25.95 AND SWEAT > 64.5 AND PRESS_UA <= 931.05 AND TTI <= 53.9	0.77
IF MONTH = 5, 6, OR 7 AND LI > -0.55 AND LI <= 2.75 AND KI > 25.95 AND SWEAT > 64.5 AND PRESS_UA > 931.05 AND SSI > -0.92 AND SSI <= 0.02	0.89
IF MONTH = 5, 6, OR 7 AND KI > 25.95 AND SWEAT > 64.5 AND PRESS_UA > 931.05 AND SSI > 0.76 AND LI > 0.65 AND LI <= 2.75	0.77
IF MONTH = 5, 6, 7, OR 8 AND KI > 21.85 AND LI > 2.75 AND PRESS_UA <= 931.5 AND SWEAT <= 190	0.75

Table B-34. Rules to predict CG lightning for Minot within 10 nm (cost two to one).

Rules to Forecast CG lightning	**Probability of correctly predicting CG lightning**
IF MONTH = 7 AND KI <= 21.85 AND SWEAT <= 75.5	0.66
IF MONTH = 5 OR 9 AND KI <= 21.85 AND LI <= 11.85 AND SWEAT > 75.5 AND SWEAT <= 146 AND TTI <= 47.7 AND SSI > -1.17 AND SSI <= 7.15	0.41
IF MONTH = 6, 7, OR 8 & KI <= 21.85 AND SWEAT > 75.5 AND SSI > -1.17 AND LI <= 11.85 AND TTI <= 38.5	0.60
IF MONTH = 6, 7, OR 8 AND KI <= 21.85 AND SWEAT > 75.5 AND SSI > -1.17 AND TTI > 38.5 AND TTI <= 50.3 AND LI <= 4.55	0.50
IF MONTH = 5, 6, 7, 8 OR 9 AND KI > 21.85 AND LI <= 3.05	0.74
IF MONTH = 5, 6, 7, OR 8 AND KI > 21.85 AND LI > 3.05 AND SWEAT <= 189.5	0.58

Table B-35. Rules to predict CG lightning for Minot within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF MONTH = 3, 4, OR 5 AND LI <= 2.75 AND KI > 28.35	0.72
IF MONTH = 6, 7, 8, or 9 AND LI <= 2.75	0.75
IF MONTH = 6, 7, OR 8 AND LI > 2.75 AND PRESS_UA <= 931.5 AND SWEAT <= 167.5	0.65
IF MONTH = 7, OR 8 AND LI > 2.75 AND PRESS_UA > 931.5 AND PRESS_UA <= 936.5 AND KO <= 5.15 AND TTI > 41.5 AND PRESC3 <= -0.1 AND CAPE > 0.5 AND SWEAT <= 114.5	0.81
IF MONTH = 6, 7, OR 8 AND LI > 2.75 AND PRESS_UA > 931.5 AND PRESS_UA <= 936.5 AND KO > 5.15	0.56

Table B-36. Rules to predict CG lightning for Minot within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF MONTH = 5, 6, 7, 8, OR 9 AND LI <= 2.35	0.74
IF MONTH = 5, 7, OR 8 AND LI > 2.35 AND LI <= 4.25 AND KI <= 19.7 AND SSI > 2.14	0.52
IF MONTH = 5, 7, 8, OR 9 AND LI > 2.35 AND LI <= 4.25 AND KI > 21.95 AND KI <= 30.35 AND SWEAT <= 167.5	0.57
IF MONTH = 6 AND LI > 2.35 AND LI <= 4.25 AND KI > 21.95 AND SWEAT <= 206	0.81
IF MONTH = 5 OR 8 AND LI > 4.25 AND KI > 9.8 AND CAPE <= 1.1 AND SWEAT > 106.5	0.56
IF MONTH = 6 OR 7 AND CAPE <= 101.2 AND LI > 4.25 AND LI <= 11.45 AND KI > 11.1 AND SWEAT <= 62.5	0.54
IF MONTH = 6 OR 7 AND CAPE <= 101.2 AND LI > 4.25 AND LI <= 11.45 AND SWEAT > 72.5 AND SWEAT <= 156.5 AND TTI <= 38.5	0.71
MONTH = 6 OR 7 AND CAPE <= 101.2 AND LI > 4.25 AND LI <= 11.45 AND SWEAT > 72.5 AND SWEAT <= 156.5 AND TTI > 39.95	0.59

Tables B-37 through B-40 are the forecast rules and probabilities for Offutt Air Force Base, Nebraska. NE Omaha/Valley (725580) was the upper air site matched with Offutt. The following CG lightning predictors were determined to produce the best success rate for Offutt: 1) Lifted Index (LI), 2) Showalters Stability Index (SSI), 3) Total Totals Index (TTI) 4) K Index (KI), 5) Month, 6) KO Index (KO), 7) Severe Weather Threat Index (SWEAT), 8) Surface pressure at Offutt matching the sounding time (STAPRESS), and 9) Convective Available Potential Energy (CAPE).

Table B-37. Rules to predict CG lightning for Offutt within 10 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF MONTH = 4, 5, 6, OR 7 AND SSI <= 4.75 AND KI <= 30.55	0.57
IF SSI <= 4.75 AND KI > 30.55	0.71

Table B-38. Rules to predict CG lightning for Offutt within 10 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF MONTH = 3, 8, OR 9 AND SSI <= 4.75 AND KI <= 30.55 AND CAPE > 77.6 AND SWEAT <= 132.5	0.65
IF MONTH = 3, 8, OR 9 AND KI <= 30.55 AND CAPE > 77.6 AND SWEAT > 132.5 AND SSI <= 0.19	0.50
IF MONTH = 4, 5, 6, OR 7 AND SSI <= 4.75 AND KI <= 30.55	0.57
IF SSI <= 4.75 AND KI > 30.55	0.71
IF MONTH = 4, 5, 6, 8, OR 9 AND SSI > 4.75 AND KI > 14.45 AND TTI <= 42.9 AND LI > 6.05 AND SWEAT > 73.5 AND KO > 10.35	0.66

Table B-39. Rules to predict CG lightning for Offutt within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF KI > 18.45 AND KI <= 33.45 AND TTI <= 57.3 AND KO <= 10.55 AND SSI <= -0.66	0.67
IF KI > 18.45 AND KI <= 33.45 AND TTI <= 57.3 AND SSI > -0.59 AND SSI <= 2.98 AND CAPE > 110.65 AND STAPRESS > 972.9 AND LI > -3.25 AND KO > -14.5 AND KO <= -1.85 AND SWEAT > 125.5 AND SWEAT <= 276.5	0.61
IF KI > 18.45 AND KI <= 33.45 AND TTI <= 57.3 AND SSI > -0.59 AND SSI <= 2.98 AND KO > -1.85 AND KO <= 10.55 AND LI <= 0.75	0.86
IF KI > 18.45 AND KI <= 33.45 AND SSI > -0.59 AND SSI <= 2.98 AND KO > -1.85 AND KO <= 10.55 AND LI > 0.75 AND TTI > 50.45 AND TTI <= 57.3	0.61
IF SSI <= 2.98 AND KI > 18.45 AND KI <= 33.45 AND TTI > 57.3	0.88
IF SSI <= 2.98 AND KI > 33.45	0.82
IF MONTH = 5, 6, 7, OR 8 AND KI > 16.65 AND KI <= 28.15 AND KO <= 11.15 AND LI <= 5.95 AND STAPRESS > 973.9 AND STAPRESS <= 977.5 AND TTI > 39.1 AND SSI > 3.54 AND SSI <= 5.66	0.74
IF SSI > 2.98 AND KI > 28.15 AND TTI <= 42.75	0.83

Table B-40. Rules to predict CG lightning for Offutt within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= -0.67 AND KI <= 30.35	0.66
IF MONTH = 3, 5, OR 6 AND KI <= 17.85 AND CAPE <= 63.3 AND KO > -5.05 AND SSI > 2.16 AND SSI <= 4.40	0.60
IF SSI > -0.67 AND SSI <= 4.40 AND KI <= 17.85 AND KO > -8.85 AND CAPE > 63.3 AND CAPE <= 256.3	0.58
IF MONTH = 4, 5, 6, 7, 8, OR 9 AND KI > 18.4 AND KI <= 30.65 AND KO <= 8.35 AND TTI > 44.55 AND SSI > -0.59 AND SSI <= 4.04 AND SWEAT <= 323.5	0.56
IF SSI <= 4.40 AND KI > 30.65	0.76
IF LI <= 3.65 AND TTI <= 41.5 AND KI > 8.25 AND KI <= 25.25 AND SSI > 4.88	0.64
IF MONTH = 4, 5, 6, 7, 8, OR 9 AND CAPE <= 0.15 AND KI > 18.65 AND SSI > 4.98 AND KO > 3.25 AND KO <= 10.35 AND LI > 3.65 AND LI <= 9.65 AND TTI > 38.2 AND TTI <= 42.1	0.64
IF MONTH = 5, 6, 8, OR 9 AND SSI > 4.40 AND KI > 8.25 AND LI > 3.65 AND KO > 10.35 AND SWEAT > 88.5	0.70

Tables B-41 through B-44 are the forecast rules and probabilities for Scott Air Force Base, Illinois. Peoria Regional Airport (725320) was the upper air site matched with Scott. The following CG lightning predictors were determined to produce the best success rate for Scott: 1) Lifted Index (LI), 2) Showalters Stability Index (SSI), 3) Total Totals Index (TTI) 4) K Index (KI), 5) Month, 6) KO Index (KO), 7) Severe Weather Threat Index (SWEAT), 8) Surface pressure at Peoria Regional Airport matching the sounding time (PRESS_UA), 9) Convective Available Potential Energy (CAPE), and 10) Three hour surface pressure change at Scott matching the sounding time (PRESC3).

Table B-41. Rules to predict CG lightning for Scott within 10 nm (cost one to one).

Rules to Forecast CG lightning	**Probability of correctly predicting CG lightning**
IF SSI <= 1.65 AND KI > 4.45 AND KI <= 26.25	0.59
IF KI > 4.45 AND KI <= 26.95 AND SSI > 1.65 AND KO <= 3.15 AND SWEAT > 62.5 AND SWEAT <= 103.5 AND LI > -0.7 AND LI <= 4.65	0.70
IF KO > 3.15 AND SWEAT <= 121.5 AND KI > 20.95 AND KI <= 26.95 AND SSI > 4.77	0.65
IF KI > 4.45 AND KI <= 26.95 AND SSI > 1.65 AND KO > 3.15 AND SWEAT > 121.5 AND TTI <= 48.2	0.56
IF KI > 26.95	0.73

Table B-42. Rules to predict CG lightning for Scott within 10 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.61 AND KI <= 28.85 AND TTI > 42.0	0.43
IF MONTH = 3, 4, 5, 6, 7, OR 8 AND SSI <= 3.61 AND KI <= 28.85 AND TTI > 42.0 AND CAPE > 47.8 AND CAPE <= 5106.7	0.59
IF SSI <= 3.61 AND KI > 28.85	0.72
IF MONTH = 4 OR 6 AND SSI > 3.61 AND LI > 7.25 AND KI > -4.2 AND KI <= 4.45 AND KO > 9.05 AND KO <= 14.9	0.60
IF MONTH = 3, 4, 5, 6, 7, OR 8 AND CAPE <= 972.3 AND PRESC3 > -1.8 AND PRESC3 <= 1.55 AND PRESS_UA > 990.05 AND SWEAT > 30.5 AND SWEAT <= 175.5 AND SSI > 3.61 AND SSI <= 15.80 AND TTI > 34.4 AND TTI <= 42.1 AND KI > 8.55 AND KI <= 19.25	0.51
IF MONTH = 3, 4, 5, 6, 7, OR 8 AND SSI > 3.61 AND PRESC3 > -1.8 AND CAPE <= 972.3 AND TTI <= 42.8 AND KI > 19.25 AND SWEAT <= 129.5	0.76
IF MONTH = 3, 4, 5, 6, 7, OR 8 AND SSI > 3.61 AND KI > 4.45 AND SWEAT <= 191.5 AND CAPE <= 972.3 AND PRESC3 > -1.8 AND PRESC3 <= 0.1 AND TTI > 44.0	0.48
IF MONTH = 3, 4, 5, 6, 7, OR 8 AND SSI > 3.61 AND KI > 4.45 AND PRESC3 > -1.8 AND SWEAT > 191.5	0.70

Table B-43. Rules to predict CG lightning for Scott within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF LI <= 5.65 AND KI > 14.6 AND KI <= 20.45 AND SSI <= 2.71 AND CAPE > 475.0 AND TTI > 42.9 AND TTI <= 47.3	0.58
IF KI <= 20.45 AND LI <= 5.65 AND SSI <= 3.67 AND TTI > 47.3 AND TTI <= 49.5	0.76
IF KI > 20.45 AND KI <= 29.05 AND SSI <= 2.04 AND SWEAT <= 250.5	0.67
IF KI > 20.45 AND KI <= 29.05 AND SSI <= 2.04 AND SWEAT > 298	0.77
IF KI > 20.45 AND KI <= 29.05 AND TTI <= 47.8 AND SSI > 2.04 AND SSI <= 4.59 AND LI <= 4.45 AND CAPE > 85.5 AND CAPE <= 227.4	0.61
IF KI > 20.45 AND KI <= 29.05 AND CAPE <= 40.4 AND LI > 5	0.62
IF KI > 29.05 AND SWEAT > 124.5 AND CAPE <= 82.45 AND TTI <= 49.5	0.79
KI > 29.05 AND SWEAT > 124.5 AND CAPE > 102.1 AND CAPE <= 252.8 AND SSI <= 1.60	0.81
KI > 29.05 AND CAPE > 263.7	0.72

Table B-44. Rules to predict CG lightning for Scott within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF KI <= 20.45 AND SSI <= 3.67 AND CAPE <= 4828.4 AND & LI <= -1.75	0.64
IF SSI > 3.67 AND TTI > 30.4 AND KI > 6.2 AND KI <= 20.45 AND LI > -1.35 AND LI <= 5.05 AND SWEAT > 67.5 AND SWEAT <= 103	0.61
IF KI > -7.35 AND KI <= 13.25 AND TTI <= 33.5 AND LI > 5.65 ANDLI <= 11.85 AND SWEAT > 53.5 AND SWEAT <= 127.5	0.53
IF SWEAT <= 193 AND KI > 13.25 AND KI <= 20.45 AND TTI <= 42.05 AND LI > 12.45	0.62
IF KI > 20.45 AND KI <= 29.05 AND SSI <= 2.04 AND TTI > 41.5 AND SWEAT > 146 AND CAPE > 53.1	0.70
IF KI > 20.45 AND KI <= 29.05 AND SSI > 2.04 AND SSI <= 4.89 AND TTI > 42.2 AND TTI <= 48.2 AND CAPE <= 51.2 AND SWEAT > 83.5 AND LI > 3 AND LI <= 11.35	0.62
IF KI > 20.45 AND KI <= 29.05 AND TTI <= 48.2 AND SSI > 2.04 AND SSI <= 4.89 AND CAPE > 85.5 AND CAPE <= 601.1 AND LI <= 3.35 AND SWEAT > 149	0.69
TTI <= 48.2 AND SSI > 4.89 AND KI > 20.45 AND KI <= 24.85 AND CAPE <= 51.1 AND LI <= 17.3	0.58
TTI <= 48.2 AND SSI > 4.89 AND KI > 24.85 AND KI <= 29.05	0.74
KI > 29.05 AND SWEAT > 124.5 AND CAPE <= 252.8	0.74
KI > 29.05 AND CAPE > 263.7	0.72

Tables B-45 through B-48 are the forecast rules and probabilities for Westover Air Force Base, Massachusetts. Albany County Airport (725180) was the upper air site matched with Westover. The following CG lightning predictors were determined to produce the best success rate for Westover: 1) Showalters Stability Index (SSI), 2) Lifted Index (LI), 3) Total Totals Index (TTI) 4) K Index (KI), 5) Convective Available Potential Energy (CAPE), and 6) Severe Weather Threat Index (SWEAT).

Table B-45. Rules to predict CG lightning for Westover within 10 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 2.58 AND LI > -1.55 AND SWEAT <= 199.5 AND KI > 20.4 AND KI <= 23.65	0.67
IF SSI <= 2.58 AND KI <= 23.65 AND LI > -1.55 AND SWEAT > 199.5	0.71
IF SSI <= 2.58 AND KI > 23.65	0.75
IF KI > 17.85 AND CAPE > 11.6 AND SWEAT > 125.5 AND SSI > 3.09 AND TTI > 37.2	0.62

Table B-46. Rules to predict CG lightning for Westover within 10 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 2.58 AND LI > -1.55 AND KI > 17.05 AND KI <= 23.65 AND SWEAT <= 183.5	0.61
IF SSI <= 2.58 AND KI <= 23.65 AND LI > -1.55 AND SWEAT > 199.5 AND CAPE > 122.0	0.78
IF SSI <= 2.58 AND KI > 23.65	0.75
IF TTI <= 49.0 AND KI > 17.95 AND SSI > 3.09 AND SSI <= 4.63 AND SWEAT > 72 AND SWEAT <= 234 AND LI <= 1.65	0.81
IF TTI <= 49.0 AND KI > 17.95 AND SSI > 3.09 AND SSI <= 4.63 AND SWEAT > 72 AND SWEAT <= 234 AND LI > 4.75 AND LI <= 6.45	0.85
IF SSI > 5.79 AND SWEAT > 77.5 AND LI <= 8.15 AND KI > 23.55 AND KI <= 25.65	0.79

Table B-47. Rules to predict CG lightning for Westover within 25 nm (cost one to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 4.43 AND KI <= 24.15 AND TTI > 45.4 AND TTI <= 50.7 AND SWEAT > 87.5	0.64
IF KI > 24.15 AND CAPE <= 182.3 AND SSI <= 2.60 AND SWEAT > 164.5	0.77
IF KI > 24.15 AND CAPE <= 182.3 AND SSI > 2.60 AND SSI <= 4.43 AND LI > 4.25 AND LI <= 7	0.68
IF SSI <= 4.43 AND KI > 24.15 AND CAPE > 182.3	0.77
IF KO > 2.85 AND SSI > 4.43 AND SSI <= 14.05 AND KI > 17.85 AND CAPE > 11.3 AND CAPE <= 64.9 AND TTI <= 39.7	0.67
IF KI > 11.55 AND CAPE > 178 AND SSI > 4.43 AND SSI <= 6.33	0.62

Table B-48. Rules to predict CG lightning for Westover within 25 nm (cost two to one).

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF KI <= 20.45 AND SSI <= 3.67 AND CAPE <= 4828.4 AND LI <= -1.75	0.61
IF SSI > 3.67 AND TTI > 30.4 AND KI > 6.2 AND KI <= 20.45 AND LI > -1.35 AND& LI <= 5.05 AND SWEAT > 67.5 AND SWEAT <= 103	0.64
IF KI > -7.35 AND KI <= 13.25 AND TTI <= 33.5 AND LI > 5.65 AND LI <= 11.85 AND SWEAT > 53.5 AND SWEAT <= 127.5	0.67
IF SWEAT <= 193 AND KI > 13.25 AND KI <= 20.45 AND TTI <= 42.05 AND LI > 12.45	0.77
IF KI > 20.45 AND KI <= 29.05 AND SSI <= 2.04 AND TTI > 41.5 AND SWEAT > 146 AND CAPE > 53.1	0.58
IF KI > 20.45 AND KI <= 29.05 AND SSI > 2.04 AND SSI <= 4.89 AND TTI > 42.2 AND TTI <= 48.2 AND CAPE <= 51.2 AND SWEAT > 83.5 AND LI > 3 AND LI <= 11.35	0.73
IF KI > 20.45 AND KI <= 29.05 AND TTI <= 48.2 AND SSI > 2.04 AND SSI <= 4.89 AND CAPE > 85.5 AND CAPE <= 601.1 AND LI <= 3.35 AND SWEAT > 149	0.62

Tables B-49 through B-52 are the forecast rules and probabilities for Wright Patterson Air Force Base, Ohio. Wilmington (724260) was the upper air site matched with WPAFB. The following CG lightning predictors were determined to produce the best success rate for WPAFB: 1) Showalters Stability Index (SSI), 2) Surface temperature at WPAFB matching the sounding time (TEMPC), 3) Surface pressure at WPAFB (STAPRESS), and 4) K Index (KI).

Table B-49. Rules to predict CG lightning for Wright Patterson within 10 nm (cost one to one)

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 1.11 AND TEMPC > 8.9	0.75
IF SSI > 1.11 AND SSI <= 3.13 AND TEMPC <= 24.7	0.58

Table B-50. Rules to predict CG lightning for Wright Patterson within 10 nm (cost two to one)

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 3.13	0.69
IF SSI > 3.13 AND STAPRESS <= 987.4 AND KI > 16.85	0.43

Table B-51. Rules to predict CG lightning for Wright Patterson within 25 nm (cost one to one)

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 0.63	0.77
IF SSI > 0.63 AND SSI <= 3.22 AND KI <= 29.35 AND STAPRESS <= 984.1	0.61
IF SSI > 0.63 AND SSI <= 3.22 AND KI > 29.35	0.69

Table B-52. Rules to predict CG lightning for Wright Patterson within 25 nm (cost two to one)

Rules to Forecast CG lightning	Probability of correctly predicting CG lightning
IF SSI <= 4.18	0.68
IF SSI > 4.18 AND KI > 12 AND STAPRESS <= 987.4 AND TEMPC > 9.7	0.44

BIBLIOGRAPHY

15th Operational Weather Squadron (OWS), 2003: 15th Operational Support Squadron, [on line at https://ows.scott.af.mil].

Air Force Weather Agency (AFWA), 1998: *Meteorological Techniques*. Technical Note TN-98/002 242 pp.

Air Weather Service (AWS), 1975: *Notes on Analysis and Severe-storm Forecasting Procedures of the Air Force Global Weather Central*., Technical Report 200, 190 pp.

Air Weather Service (AWS), 1990: *The Use of the Skew-T, Log P Diagram in Analysis and Forecasting*, Technical Reference 79-006, 153pp.

Allen, M., 2003: Evaluation of the Mountain Wave Forecast Model's Stratospheric Turbulence Simulations. M.S. Thesis, Department of Engineering Physics, Air Force Institute of Technology, 81 pp.

AMS, 2002: *A Brief Guide for Authors.* American Meteorological Society, Boston, Massachusetts.

Benz, R.F., 2003: Data Mining Atmospheric/Oceanic Parameters in the Design of a Long-Range Nephelometric Forecast Tool. M.S. Thesis, Department of Engineering Physics, Air Force Institute of Technology, 88 pp.

Bidner, A., 1970: The Air Force Global Weather Central Severe Weather Threat (SWEAT) Index – A Preliminary Report. *Air Weather Service Aerospace Science Review*, AWS RP 105-2, no. 70-3, 2-5.

Blanchard, D. O., 1998: Assessing the vertical distribution of Convective Available Potential Energy. Wea. Forecasting, **13**, 870-877.

Bluestein, H.B., 1992: *Synoptic-Dynamic Metereorology in Midlatitudes Volume I*. Oxford University Press, 431 pp.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J., 1984: *Classification and Regression Trees*. Wadsworth, Belmont, CA, 358 pp.

Cummins K.L., M. J. Murphy, E.A. Bardo, W. L. Hiscox, R. B. Pyle, and A. E. Pifer, 1998: A Combined TOA/MDF Technology Upgrade of the U.S. National Lightning Detection Network. *J. Geophys. Res*., **103**, 9035-9044.

Freestrom, H.J., 2001: Designing an Algorithm to Predict the Intensity of the Severe Weather Season. M.S. Thesis, Department of Engineering Physics, Air Force Institute of Technology, 71 pp.

Galway, J.G., 1956: The Lifted Index as a Predictor of Latent Instability. Bull. Amer. Meteor. Soc., **37**, 528-529.

George, J.J., 1960: *Weather Forecasting for Aeronautics*. Academic Press, 673 pp.

GHRC, 2003: An Introduction to the National Lightning Detection Network, [on line at http://ghrc.msfc.nasa.gov:5721/dataset_documents/NLDN_dataset_doc.html].

Golden J.H., R. Serafin, V. Lally, J. Facundo, 1986: Atmospheric Sounding Systems. *Mesoscale Meteorology and Forecasting*, P.S. Ray, Ed., American Meteorological Society, 50-70.

Holton, J.R., 1992: *An Introduction to Dynamic Meteorology*. Academic Press, 511 pp.

Huffines, G.R., and R.E. Orville, 1999: Lightning Ground Flash Density and Thunderstorm Duration in the Continental United States. *J. Appl. Meteor.,* **38**, 1013-1019.

Huschke, R.E., 1959: *Glossary of Meteorology*. American Meteorological Society, 638 pp.

Lewis, Roger J., 2000: *An Introduction to Classification and Regression Tree (CART) Analysis*. Department of Emergency Medicine, Harbor-UCLA Medical Center, Torrance, CA.

Lightning Storm; 2003: National Lightning Detection Network. [online at http://www.lightningstorm.com/tux/jsp/discover/nldn/index.jsp;jsessionid=70302723151069687033662].

Lyons W.A., Uliasz M., Nelson T. E., 1998: Large Peak Current Cloud-to-Ground Lightning Flashes during the Summer Months in the Contiguous United States. *Mon. Wea. Rev.*, **126**, 2217-2233.

Miller, R.C.,1972: *Notes on analysis and severe storms forecasting procedures of the Air Force Global Weather Central*. Tech. Rep. 200 (rev.), Air Weather Service.

Miller, R.C., A. Bidner, and R.A. Maddox, 1971: *The Use of Computer Products in severe weather forecasting (the SWEAT index)*. Seventh Conference Severe Local Storms, Kansas City Amer. Meteor. Soc., 1-6.

Montgomery, D.C. and G.C. Runger, 2003: *Applied Statistics and Probability for Engineers Third Edition*. John Wiley and Sons, 706 pp.

Mueller, C. K., J. W. Wilson, and N. A. Crook, 1993: The utility of sounding and mesonet data to nowcast thunderstorm initiation," *Wea. Forecasting*, **8**, 132-146.

National Lightning Safety Institute, 2003: Lightning Incidents, [online at http://www.lightningsafety.com/nlsi_lls.html].

National Oceanic and Atmospheric Administration (NOAA), 1997: *Lightning Fatalities, Injuries, and Damage Reports in the US from 1959-1994*. NOAA Technical Memorandum NWS SR-193.

National Weather Service (NWS) Louisville 2003: Convective Parameters and Indices, [online at http://submit.crh.noaa.gov/lmk/soo/docu/indices.htm].

Orville, R.E., and G.R. Huffines, 1999: Lightning Ground Flash Measurements over the Contiguous United States," *Mon. Wea. Re*v., **127**, 2693-2703.

Orville, R.E., and G.R. Huffines, 2001: Cloud-to-Ground Lightning in the United States: NLDN Results in the First Decade, 1989-98., *Mon. Wea. Re*v., **129**, 1179-1193.

Orville, R.E., and G.R. Huffines, W.R. Burrows, R.L. Holle, and K.L. Cummins, 2002: The North American Lightning Detection Network (NALDN) – First Results: 1998-2000, *Mon. Wea. Re*v., **130**, 2098-2109.

Peppler, Randy A., 1998: A Review of Static Stability Indices and Related Thermodynamic Parameters. SWS Miscellaneous Pub. 104, Illinois State Water Survey Division, October 1988, 87 pp.

Randall, R.M., 2002: Exploration of Teleconnection Indices for Long-Range Seasonal Temperature Forecasts. M.S. Thesis, Department of Engineering Physics, Air Force Institute of Technology, 88pp.

Rauber, R.M., J.E. Walsh, and D.J. Charlevoix, 2002: Severe and Hazardous Weather. Kendall/Hunt Publishing Company, 616 pp.

Reap, R. M.,1994: Analysis and Prediction of Lightning Strike Distributions Associated with Synoptic Map Types over Florida. *Mon. Wea. Rev.*,**122,** 1698-1715.

Sahu, J. K., and S. V. Singh, 1999: Single Station Forecasting of Thunderstorms From Large Scale Features. *Advanced Technologies in Meteorology*, Ch 1 [on line at http://www.commonwealthknowledge.net/MetCD/Chapter1/C1P02.htm].

Salford Systems, 1995: *Data Mining with Decision Trees: An Introduction to CART®*. San Diego, CA, 133 pp.

Salford Systems, 2002: *CART® for Windows User's Guide*. San Diego, CA, 295 pp.

Showalter, A.K., 1953: Stability Index for Forecasting Thunderstorms," *Bull. Amer. Met. Soc*., **34**(6), 250-252.

South African Weather Bureau, 2000: Background Information on instability indices and parameters related to Thunderstorm Development. [online at http://www.weathersa.co.za/wfr/fcastaids/pcgrids/tsback.htm].

Venske, K. C., 2001: Developing Predictors for Cloud-to-Ground Lightning Activity using Atmospheric Stability Indices. M.S. Thesis, Department of Engineering Physics, Air Force Institute of Technology, 186 pp.

Wacker, R. S., and R. E. Orville, 1999a: Changes in measured lightning flash count and return stroke peak current after the 1994 U. S. National Lightning Detection Network upgrade. Part I: Observations. *J. Geophys. Res*., **104**, (D2), 2151-2157.

Wacker, R. S., and R. E. Orville, 1999b: Changes in measured lightning flash count and return stroke peak current after the 1994 U. S. National Lightning Detection Network upgrade. Part II: Theory. *J. Geophys. Res*., **104**, (D2), 2159-2162.

Wallace J.M. and P.V.Hobbs, 1977: *Atmospheric Science an Introductory Survey*. Academic Press, 467 pp.

Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Vita

Captain Manuel I Folsom Jr. graduated from Lafayette High School in Mayo, Florida. He completed his Associate Degree from North Florida Junior College. Later that year, he enlisted in the Air Force. He returned to school to complete his undergraduate studies at Florida State University in Tallahassee, Florida where he graduated with a Bachelor of Science degree in Meteorology. He was commissioned through Officer Training School at Maxwell AFB in Montgomery, Alabama.

His first assignment was at Little Rock AFB as a life support technician. While there, he was accepted into the Airman Education and Commissioning Program and was reassigned to Florida State University to complete his undergraduate degree. His first assignment after commissioning was to Barksdale AFB in Shreveport Louisiana. He served as the wing weather officer for 21 months. His next assignment was to Seoul, Korea for a one year short tour assignment. Finally, he entered the Graduate School of Engineering and Management, Air Force Institute of Technology. Upon graduation, he will be assigned to the Air Force Combat Climatology Center in Asheville, North Carolina.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
Mar 2004	Master's Thesis	Jun 2003 – Mar 2004

4. TITLE AND SUBTITLE

DEVELOPING A FORECAST TOOL FOR CLOUD TO GROUND LIGHTNING IN THE NORTH CENTRAL AND NORTHEASTERN UNITED STATES

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Folsom., Manuel I Jr. Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/ENP)
2950 Hobson Way, Bldg 640
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GM/ENP/04-05

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

15th Operational Weather Squadron
Lt Col Zuccarello
102 West Losey Street
Scott AFB, IL 62225

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Cloud-to-ground (CG) lightning is a hazard to the Air Force for both air and ground operations. Forecasting CG lightning is a necessary and extremely important requirement for Air Force meteorologists and forecasters. The 15th Operational Weather Squadron requested a forecast tool capable of predicting CG lightning within a 25 and 10 nautical mile radius of the 13 military locations in their area of responsibility. To fulfill their request, forecast decision tools were created using a Classification and Regression Tree (CART) data analysis program. Four decision trees were produced for each location using the period of record from March through September, 1993 to 2002. CART compared the upper air stability indices and surface data at 12-hour intervals with CG lightning data occurring within the next 12 hours to determine prediction rules. Data from 2003 were used as independent verification of the decision trees. The CART decisions trees were examined using contingency tables and verification tests to determine the value of the products created. The straight forward forecast rules and the verification test results confirmed that the decision trees would be a valuable prediction tool. Combined with forecaster knowledge, forecast models and other tools, the decision trees would provide an excellent forecast method for determining the occurrence of CG lightning. Therefore, the results are recommended to the 15th Operational Weather Squadron for use as a CG lightning forecast tool.

15. SUBJECT TERMS

Cloud to ground Lightning, Stability Indices

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Lt Col Michael K. Walters, ENP
U	U	U	UU	134	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-7419, ext 4642

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18